INVESTIGACION Y CIENCIA

Edición española de **SCIENTIFIC AMERICAN**



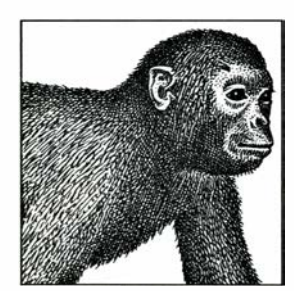
PAUTAS DE COMPLEJIDAD

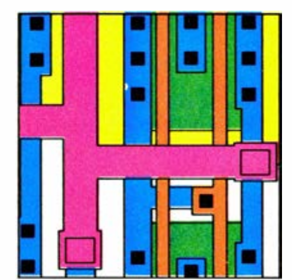
Marzo 1989
500 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

INVESTIGACION Y
CIENCIA
Marzo de 1989 Número 150

Edición española de **SCIENTIFIC AMERICAN**







SECCIONES



COLABORADORES DE ESTE NUMERO

Asesoramiento y traducción:

Santiago Torres Martínez: *Activadores génicos*; Montserrat Domingo Morató: *Terremotos profundos*; Ramón Pascual: *Mezcla de fluidos*; Ana García: *Hidratos de carbono y depresión*; Elisa Villa Otero: *La caza de* Proconsul; Luis Bou: *El problema de la red mínima* y *Juegos de ordenador*; José Chabás: *André-Marie Ampère*; J. Vilardell: *Taller y laboratorio*.

Ciencia y sociedad:

Josep-Enric Llebot

Ciencia y empresa:

Manuel Puigcerver

Libros:

Jordi Agustí, Claudi M. Cuchillo y Luis Alonso

PROCEDENCIA DE LAS ILUSTRACIONES

Página	Fuente
8	Eladio Jaraíz
10-16	Eladio Jaraíz y Magda Mària
19	Janice A. Fischer, Universidad de California, Berkeley
20-25	Neil O. Hardy
27	Neculai Mandrescu, Instituto de Geofísica, Bucarest
28-31	Joe Lertola
32	Joe Lertola (*arriba a la izquierda*), William A. Bassett (*arriba a la derecha*), Joe Lertola (*abajo*)
33	Joe Lertola
45-47	C. W. Leong y Julio M. Ottino, Universidad de Massachusetts en Amherst
48	Bob Conrad
49	Bob Conrad (*arriba*), C. W. Leong y Julio M. Ottino (*abajo*)
50	Bob Conrad (*arriba*); John G. Franjione y Julio M. Ottino (*abajo*)
51	Paul D. Swanson y Julio M. Ottino, Universidad de Massachusetts en Amherst
54	K. R. Sreenivasan, Universidad de Yale
55	Ichiro Sugioka y Bradford Sturtevant, Instituto de Tecnología de California
57	Robert Mankoff, © 1984 por *The New Yorker Magazine*, Inc.
58	Andrew Christie
59	Laurie Burnham
60	Andrew Christie
61	S. Varnedoe
62-63	Andrew Christie
67	Alan Walker
68-72	Tom Prentiss
75	Quesada/Burke (*arriba*), Gabor Kiss (*abajo*)
76-79	Gabor Kiss
83	The Granger Collection
84	L. Pearce Williams
85-88	Hank Iken
90-93	Michael Goodman
95-96	Andrew Christie
98	Johnny Johnson



LA FOTOGRAFIA DE LA PORTADA muestra la compleja estructura que una pequeña gota de trazador fluorescente puede producir en un fluido viscoso si la gota se coloca en una zona de mezcla caótica (véase "Mezcla de fluidos", por Julio M. Ottino, página 44). El flujo mezclador se generó mediante un movimiento periódico de contrarrotación de dos cilindros excéntricos. Aunque tal flujo es caótico, tiene ciertas simetrías. Se muestra una "isla" de fluido sin mezclar rompiéndose en dos islas más pequeñas al pasar por el eje de simetría.

INVESTIGACION Y CIENCIA

DIRECTOR GENERAL Francisco Gracia Guillén
DIRECTOR EDITORIAL José María Valderas Gallardo
PRODUCCIÓN César Redondo Zayas
PROMOCIÓN EXTERIOR Pedro Clotas Cierco
EDITA Prensa Científica, S. A. Calabria, 235-239 - 08029 Barcelona (ESPAÑA)
Teléfono 322 05 51 Telefax 410 79 12

SCIENTIFIC AMERICAN

EDITOR Jonathan Piel
BOARD OF EDITORS Armand Schwab, Jr., Managing Editor; Timothy Appenzeller, Associate Editor; Timothy M. Beardsley; John M. Benditt; Laurie Burnham; Elizabeth Corcoran; Ari W. Epstein; Gregory R. Greenwell; John Horgan; June Kinoshita; Philip Morrison, Book Editor; Tony Rothman; Ricki L. Rusting; Russell Ruthen, Karen Wright
PUBLISHER Harry Myers
ADVERTISING DIRECTOR Peter B. Kennedy
PRESIDENT AND CHIEF EXECUTIVE OFFICER Claus-Gerhard Firchow
CHAIRMAN OF THE BOARD Georg-Dieter von Holtzbrinck
CHAIRMAN EMERITUS Gerard Piel

SUSCRIPCIONES

Prensa Científica S. A.
Calabria, 235-239
08029 Barcelona (España)
Teléfono 322 05 51 ext. 33-37

Precios de suscripción, en pesetas:

	Un año	Dos años
España	5500	10.000
Extranjero	6200	11.500

Ejemplares sueltos:

Ordinario: 500 pesetas
Extraordinario: 650 pesetas

- Todos los precios indicados incluyen el IVA, cuando es aplicable.
- En Canarias, Ceuta y Melilla los precios incluyen el transporte aéreo.
- El precio de los ejemplares atrasados es el mismo que el de los actuales.

DISTRIBUCION

para España:

MIDESA
Carretera de Irún, km. 13,350
(Variante de Fuencarral)
28049 Madrid Tel.652 42 00

para los restantes países:

Editorial Labor, S. A.
Calabria, 235-239 - 08029 Barcelona
Teléfono 322 05 51

PUBLICIDAD

Madrid: Gustavo Martínez Ovín
Menorca, 8, bajo, centro, izquierda.
28009 Madrid
Teléfonos 409 70 45 - 409 70 46
Cataluña:
Thaïs Muñoz Oliva
Muntaner, 467-469, 3° 3ª
08021 Barcelona
Teléfono 211 92 45



ISSN 0210-136X Dep. legal: B. 38.999-76

Fotocomposición: Tecfa. Línea Fotocomposición, S.A. Almogàvers, 189 - 08018 Barcelona

Fotocromos reproducidos por Reprostudy, S. A. Buenos Aires, 51 - 08902 Hospitalet (Barcelona)

Gráfica Elzeviriana, S. A. Nápoles, 249 - Tel. 207 40 11 - 08013 Barcelona

Printed in Spain - Impreso en España

Hace...

José M.ª López Piñero

...*trescientos cincuenta años*

Juan Caramuel publicó *Coelestes Metamorphoses*, primera de sus obras de astronomía.

Nacido en Madrid de ascendencia bohemia y flamenca, Juan Caramuel y Lobkowitz (1606-1682) es una figura que corresponde plenamente al carácter exuberante y poco sereno que tópicamente suele asociarse al período barroco. Después de cursar humanidades en Alcalá, tomó el hábito cisterciense. Completó más tarde su formación en diferentes centros españoles y extranjeros, doctorándose por último en teología en la Universidad de Lovaina. Protegido por el monarca español Felipe IV, por el emperador Fernando III y por el papa Alejandro VII, desempeñó importantes cargos eclesiásticos, hasta ser nombrado obispo de Campagna, en el reino de Nápoles, y luego de Vigevano, en el Milanesado. La mayor parte de su vida la pasó en Bohemia y en distintas localidades de los dominios españoles de Italia. Esto le permitió una independencia y un conocimiento de las nuevas corrientes europeas excepcionales entre los intelectuales y científicos residentes en la misma España. No obstante, su ausencia desde los años de juventud le impidió ejercer una influencia directa en el desarrollo de nuestra ciencia. A pesar de su continuada conexión con los problemas del país y del "españolismo" de su obra, puesto de relieve por Ceñal, ni su biografía ni su producción científica pueden considerarse con propiedad integradas en la realidad social española de su época. Interesa más bien como un "español fuera de España" cuya posición de adelantado y cuyo prestigio fueron después indirectamente aprovechados por los *novatores* residentes en el país.

Caramuel no fue un profesional de la ciencia, ni siquiera un clérigo consagrado preferentemente a los estudios científicos, como lo fueron Zaragoza o Tosca. Su producción escrita es famosa ante todo por el extraordinario número de títulos que comprende y por la diversidad de materias sobre las que trata. Buena parte está dedicada a la teología y a variados temas religiosos o eclesiásticos. Incluye asimismo estudios de musicología, de gramática y de estética, textos de carácter político e histórico, además de obras de contenido filosófico y una serie nada escasa de libros científicos. No hace falta subrayar que tan desmesurada fecundidad perjudicó inevitablemente el rigor de sus exposiciones. Tampoco favoreció este último su afán de originalidad, que le llevó en ocasiones a defender auténticas extravagancias o falsas novedades basadas en débiles argumentos.

La importancia de la obra científica de Caramuel queda desenfocada si atendemos a los encomiásticos pero interesados elogios que de ella hicieron los innovadores españoles del período siguiente, como Zapata, Sarmiento o Feijoo, que todavía resuenan en los juicios de Menéndez Pelayo. Su interés, en efecto, no reside en supuestas aportaciones creadoras, ni siquiera en su solidez, sino en que estuvo ampliamente abierta a las nuevas tendencias en unas fechas relativamente tempranas. Fue de esta manera uno de los primeros puentes de comunicación entre las corrientes modernas europeas y el ambiente científico español en lo relativo a cuestiones matemáticas, astronómicas y físicas.

Su labor en estas disciplinas dependió estrechamente de Descartes. Ceñal, que ha puesto muy agudamente de relieve esta dependencia en la esfera del pensamiento filosófico, no se ha limitado a recoger su relación directa con el gran filósofo y científico francés, sino que ha subrayado su pertenencia a una "órbita cartesiana" de autores europeos, entre los que se encuentran Gassendi, Maignan y la compleja figura de Atanasio Kircher, de personalidad tan afín a la suya.

La física de Caramuel continúa siendo filosofía natural, aunque en abierta ruptura con los principios clásicos. Por ello la expone principalmente en sus escritos filosóficos, en los que aparece como un adversario decidido de las doctrinas aristotélicas. "Instituto novam accademiam contra peripateticam" afirma, por ejemplo, en la dedicatoria de su *Rationalis et realis Philosophia* (1642). Se opone de este modo al hilemorfismo, sumándose en varias obras a la teoría cartesiana de las cualidades de los cuerpos y de los modos de la extensión y también a algunos puntos de vista de Gassendi. Esta línea no es, como vemos, la de la auténtica física moderna, pues continuaba siendo tan apriorista y especulativa como la tradicional. Como la de su modelo, Descartes, demostraba una clara simpatía por las nuevas corrientes, principalmente por su tendencia al mecanicismo y al matematismo, pero no formaba parte todavía de una disciplina positiva e independizada de la filosofía. Entre la abundante producción de Caramuel es posible encontrar algún escrito de orientación más propiamente física, como *Sublimium ingeniorum crux* (1642), dedicado al problema de la caída de los graves sobre base experimental. También incluyó numerosos problemas físicos junto a otros filosóficos y teológicos en su *Mathesis audax* (1642), pero no llegó nunca a ofrecer una exposición sistemática de esta disciplina dentro de un marco filosófico sino matemático, como después hizo Tosca.

Más interesante desde el punto de vista puramente científico es su producción astronómica. Aunque no fue un astrónomo práctico como Vicente Mut o José de Zaragoza, se le deben interesantes exposiciones teóricas, contenidas en sus libros filosóficos o científicos y también en algunos estudios monográficos. El primero fue, como hemos adelantado, el titulado *Coelestes Metamorphoses* (1639) o transforma-

NUMERO DE DADOS		ESPERANZA	PELIGRO	POR LO TANTO, PAGAS	POR LO TANTO, RECIBES
2	12	1	35	1	35
3	11	2	34	1	17
4	10	3	33	1	11
5	9	4	32	1	8
6	8	5	31	1	$6\frac{1}{5}$
7	7	6	30	1	5

1. Tabla de probabilidades en juegos de azar sacada del Cursus mathematicus *de Caramuel*

ciones de las órbitas circulares teóricas de los planetas. "No es pequeña gloria para Caramuel —afirma Cotarelo de este texto— hacerse cargo, aunque sea para rechazarla, de la primera ley de Kepler, generalmente inadvertida entonces." En otro estudio, *De novem sideribus circa Jovem visis* (1643), defendió al capuchino Reitha que creyó haber descubierto cinco satélites nuevos en torno a Júpiter. Pero la más importante de sus aportaciones en este terreno es la contenida en la cuarta parte de su *Cursus Mathematicus*, que lleva el expresivo título de *Mathesis Astronomica in Physicae Tribunali damnata* (1668). Lo que no había llegado a hacer con la física lo hizo con la astronomía, que aparece aquí como una ciencia positiva formulada en lenguaje matemático, es decir, en la auténtica línea de la "mecánica celeste".

La parte más valiosa de la obra científica de Caramuel es la matemática. Su obra fundamental en este terreno es el *Cursus Mathematicus* (1667-1668), ambicioso tratado en cuatro partes, de las cuales las dos primeras —*Mathesis Vetus novis operationum compendiis et demonstrationibus dilucidata* y *Mathesis Nova, veterum inventis confirmata*— son de contenido matemático, la segunda —*Mathesis Arquitectonica*— está dedicada a la arquitectura y a la técnica de la construcción, y la tercera, como ya hemos dicho, a la astronomía. Aparecida en 1667-1668 la primera edición, en 1670 se publicó una segunda de las dos primeras partes, con el título de *Mathesis biceps, vetus et nova*. También de la tercera parte apareció otra edición, traducida al castellano por el propio Caramuel, que la tituló *Architectura civil recta, y obliqua* (1678). Resulta muy significativo que la dedicara a Juan José de Austria.

Garma destaca, entre las contribuciones matemáticas de Caramuel, su estudio completo de los sistemas de numeración, la exposición de la teoría del número entero en forma analítica, el estudio detallado y moderno de la combinatoria, y el planteamiento y resolución de los problemas planteados inicialmente por el cálculo de probabilidades. Publicó también las primeras tablas de logaritmos editadas por un científico español e ideó un sistema que denominó de "logaritmos perfectos". Aunque ya Zaragoza demostró que no representaban ninguna ventaja sobre los logaritmos de Briggs, esta invención coloca a Caramuel, en opinión de Sánchez Pérez, entre los antecesores del cálculo de los cologaritmos.

Como ejemplos de la producción científica de Caramuel vamos a ofrecer

Numeri naturales.	Logarithmi Vlacquii.	Logarith. Neperi, & Kepleri.	Logar. Caramuelis.
1	0.	2302,585.20.	10.
10	1.	2072,326.68.	9.
100	2.	1842,068.16.	8.
1,000	3.	1611,809.64.	7.
10,000	4.	1381,551.12.	6.
100,000	5.	1151,292.60.	5.
1;000,000	6.	921,034.08.	4.
10;000,000	7.	690,775.56.	3.
100;000,000	8.	460,517.04.	2.
1,000;000,000	9.	230,258.52.	1.
10,000;000,000	10.	000,000.00.	0.

2. Comparación entre los logaritmos de Ulacq, Neper y los "perfectos" de Caramuel, según el Cursus mathematicus *(1667-1668) de este último*

dos fragmentos de sus obras. El primero pertenece al *Cursus mathematicus* y se refiere a la matematización de los juegos de azar.

"Si jugamos con dados comunes, que tienen seis lados, y alguien apuesta una moneda de oro, diciendo que *saldrá un seis y ningún otro número*, si no sale un seis, deberá pagar una moneda de oro, y si sale, recibir cinco. Esto es válido si usamos un solo dado.

"¿Qué sucede si usamos dos o tres dados? Si alguien con dos dados dice que *van a salir 12 puntos*, tiene un solo modo de acertar y 35 de equivocarse. Por lo tanto, si pone en peligro una moneda de oro, debe esperar 35. Considera la tabla siguiente... [*figura 1*].

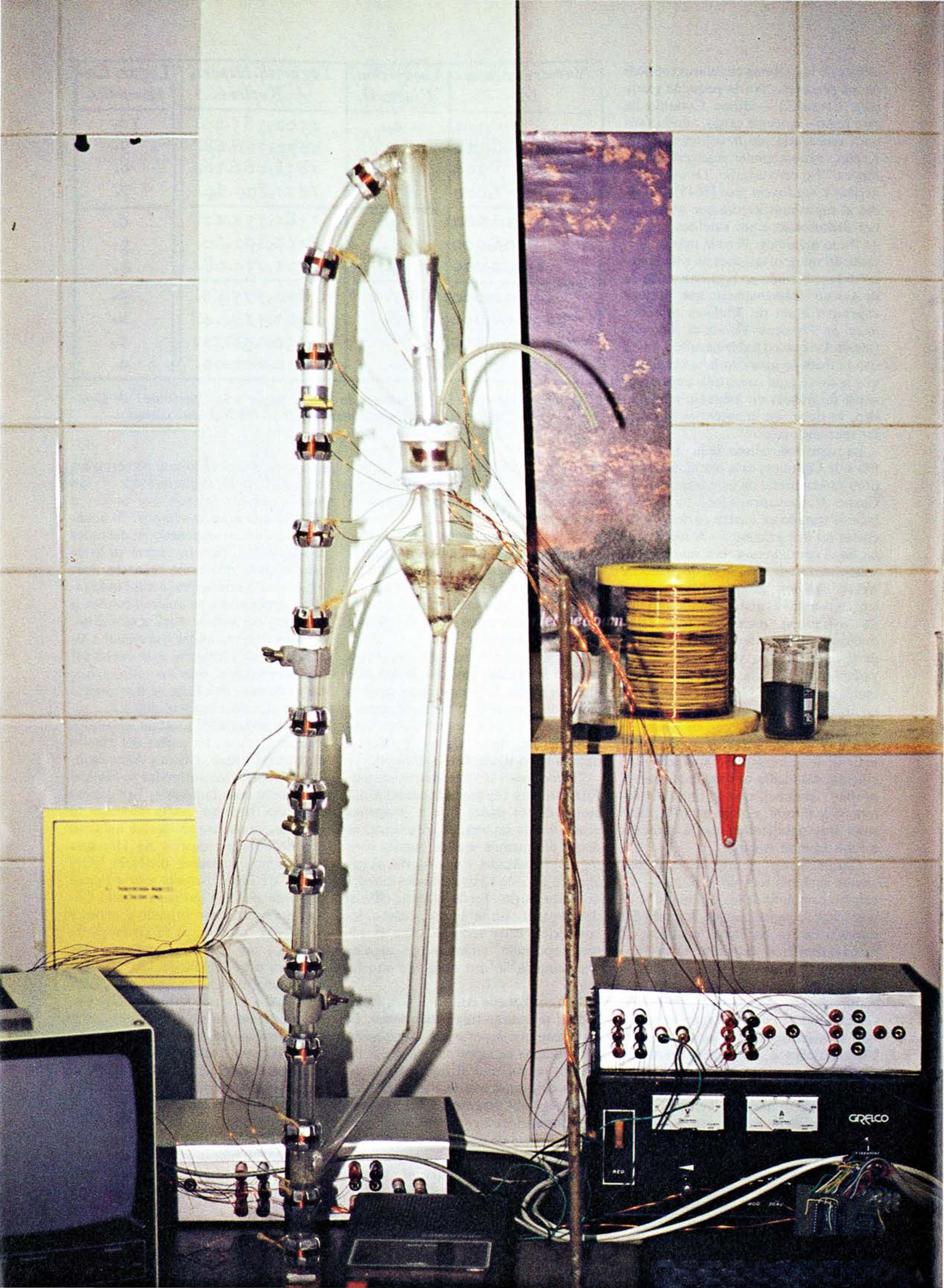
"Tienes ante la vista seis columnas. En la primera y la segunda están los números de los dados que te propones echar. En la tercera, la esperanza, es decir, de cuántos modos puedes conseguir lo apostado. En la cuarta, el peligro, o sea, de cuántos modos puedes no conseguirlo. En la quinta, cuánto debes pagar. En la última, cuánto debes recibir cada vez que salga el número apostado, mientras que pierdes una moneda de oro cada vez que no salga.

"De aquí puede deducirse lo que sucede si se utilizan tres o más dados. El que tiene que esperar una puntuación determinada siempre se expone a un peligro mayor que el que apuesta de forma disyuntiva por una u otra."

El segundo fragmento, procedente de la *Architectura civil recta y obliqua*, refleja la idea de progreso científico de Caramuel, al hablar de los logaritmos.

"Es cuestión muy reñida, promovida más por prejuicios de ingenios pusilánimes que gobernada por razón, la que pregunta: ¿Qué es lo que supieron los antiguos? Si lo supieron todo y tan todo, ¿qué no dejaron que buscar o cosa de más a los modernos? Si aquéllos se hubiesen empeñado en defender la afirmación, tuviera excusa su arrogancia; pero que éstos sean tan enemigos de sí mismos que, contra toda razón y experiencia, se quieran condenar y abatir, me parece cosa maravillosa. No es mi intento quitar o cercenar a los antiguos las alabanzas que merecen; supieron mucho, mas no todo, y hoy los modernos, con haberse descubierto nuevas ciencias, siempre dejan que saber a la posteridad. Estaban sosegados los cosmógrafos, las partes del mundo eran tres: Europa, Africa y Asia; cuanto se decía de los antípodas se oía con risa, como cosa imposible. Era común sentencia que no podía navegar al Océano, ya que no obedecían nuestras expediciones al imperio de Hércules que las puso término, diciendo NON PLUS ULTRA, *No hay más a donde se puede andar o navegar*... Salió Colón, navegó y, descubriendo grandes y ricas islas, probó que la cosmografía antigua... nos había encarcelado en los términos del NON PLUS ULTRA que en España puso la Antigüedad. PLUS ULTRA, a pesar de Hércules, dijo Colón y salió con la suya...

"Del mismo modo sucedió en el orbe literario... en nuestro tiempo los que profesan letras con generosa audacia derribaron las columnas que habían erigido los Hércules antiguos y borraron la inscripción que decía NON PLUS ULTRA, *No hay ya más que saber* y, engolfándose en el Océano de la Matemática, descubrieron una India ultramarina que hoy se llama *Logarítmica*."


GRELCO

Recirculación magnética de sólidos

Muchos procesos industriales necesitan la circulación de partículas sólidas en circuito cerrado. En el laboratorio del autor se ha conseguido el control de ese movimiento por campos magnéticos

Eladio Jaraíz

Hay procesos industriales que precisan la circulación de partículas sólidas en circuito cerrado. Imagínese el lector un reactor donde se esté desarrollando una reacción fuertemente térmica, y necesite, por tanto, aportación externa de calor, o se produzca una reacción fuertemente endotérmica, y haya que retirarlo. Recuerde, a este propósito, que los gases tienen una capacidad calorífica relativamente pequeña, si la comparamos con los calores de reacción; por contra, la capacidad calorífica de los sólidos es notable. En virtud de ello, la circulación de partículas sólidas entre un reactor y un cambiador de calor, dos recipientes, puede transportar grandes cantidades de calor de un recipiente a otro, así como establecer un control adecuado de la temperatura en el reactor y en el cambiador. En el caso en que se empleasen gases para retirar o aportar calor, esa circulación no sería viable, ni el régimen dominante en el interior del reactor sería isotérmico.

La circulación y el control de las partículas se efectúa por medio de corrientes de gases auxiliares y por válvulas, que requieren la presencia de soplantes o compresores, por motores y por otras partes móviles que exigen un mantenimiento frecuente y cuidadoso. En lo concerniente a la flexibilidad del sistema, se ha de reconocer su limitación, que no permite variar grandemente las condiciones del transporte neumático, si no se quieren tener serios problemas de control. (Se habla aquí de transporte neumático porque las partículas son arrastradas por un gas.)

1. RECIRCULADOR MAGNETICO de sólidos, desarrollado en 1986 en el departamento de ingeniería química de la Universidad de Salamanca por G. Martín y el autor. Aunque este modelo pionero no tenía las bobinas optimizadas pudo recircular partículas de acero inoxidable de tipo ferromagnético. Para el transporte ascensional se empleó un tubo de metacrilato de 20 milímetros de diámetro, a cuyo alrededor había 14 bobinas iguales con una separación entre las mismas de 50 mm. El sistema descendente consta de un tubo de caída libre para las partículas hasta llegar a la válvula magnética de sólidos que regula el flujo de partículas de retorno.

Ante esa situación, se me ocurrió en 1983 intentar demostrar la posibilidad de utilizar sólo campos magnéticos variables para efectuar el control y la circulación de partículas. Este sistema tendría muchas ventajas sobre el transporte neumático: no se necesitaría, por ejemplo, ninguna corriente de gas auxiliar para el transporte de partículas, ni soplantes, compresores o motores para que circule el gas auxiliar, lo que redundaría en un menor coste de mantenimiento de las instalaciones. Cabría también la creación de sistemas enteramente cerrados de transporte de calor por medio de sólidos que, en el ámbito de la industria nuclear o de residuos tóxicos, revisten un obvio interés general.

Los efectos de los campos magnéticos se conocen desde antiguo, desde que se tuvo relación con la magnetita, el imán permanente que se encuentra en la naturaleza. El descubrimiento de la propiedad de la misma de orientarse hacia polos antagónicos, norte y sur, ejerció un beneficioso influjo en la navegación y exploración geográficas. Pero, salvo por esa aplicación, se hizo muy poco uso del magnetismo y no hubo ningún adelanto teórico sobre la naturaleza del mismo. Hubo que esperar hasta el siglo XIX para que Hans Christian Oersted descubriera que la corriente eléctrica producía un campo magnético, y se empezara a trabajar en sus aplicaciones. En esa línea, Karl Friedrich Gauss, Joseph Henry, Michael Faraday, James Clerk Maxwell y otros mostraron la íntima relación existente entre el campo eléctrico y el campo magnético. Gracias a su esfuerzo fueron apareciendo motores, transformadores y demás ingenios donde intervienen fenómenos magnéticos; de muy reciente creación es la resonancia magnética nuclear, que posibilita la determinación de la estructura de las moléculas.

El caso más sencillo de producción artificial de campos magnéticos es el que se genera en una espira circular de un conductor por donde circula una corriente eléctrica de determinada intensidad. La corriente crea a su alrededor un campo magnético cuya magnitud viene determinada por la intensidad del flujo eléctrico y la distancia y posición del punto donde se esté calculando dicho campo magnético. El campo es una magnitud vectorial, dotada por tanto de dirección y sentido.

Algo más complejo, pero de mayor interés técnico también, es el caso de un solenoide o bobina, formado por espiras uniformemente arrolladas hasta configurar un cilindro de determinado radio y longitud. La inducción magnética, así se llama al producto de la permeabilidad magnética del medio por el valor de la intensidad del campo, se puede obtener en cualquier punto; la inducción es proporcional al número de espiras, a la intensidad de corriente que circula por el conductor que forma la bobina y a un factor geométrico que re-

ELADIO JARAIZ es profesor de ingeniería química en la Universidad de Salamanca, donde se graduó y obtuvo el doctorado con el estudio de un nuevo modelo de distribución de luz en un reactor tubular fotoquímico que tiene en cuenta la refracción de la luz; ese modelo se aplica a la fotooxidación de la anilina en disolución acuosa. Pasó dos años largos en el departamento de ingeniería química de la Universidad estatal de Oregon. De nuevo en la universidad salmantina, desde 1983 viene investigando en el desarrollo de equipos basados en la aplicación de los campos magnéticos, objeto de este artículo.

laciona la situación del punto donde se está calculando la inducción magnética con respecto a las espiras que integran la bobina. Este cálculo, cuando se trata de puntos situados sobre el eje de la bobina, resulta bastante sencillo; pero cuando se trata de puntos alojados fuera de dicho eje la solución, más compleja, ha de recurrir a estructuras matemática más refinadas: integrales elípticas de primer y segundo grado o valores tabulados de las mismas.

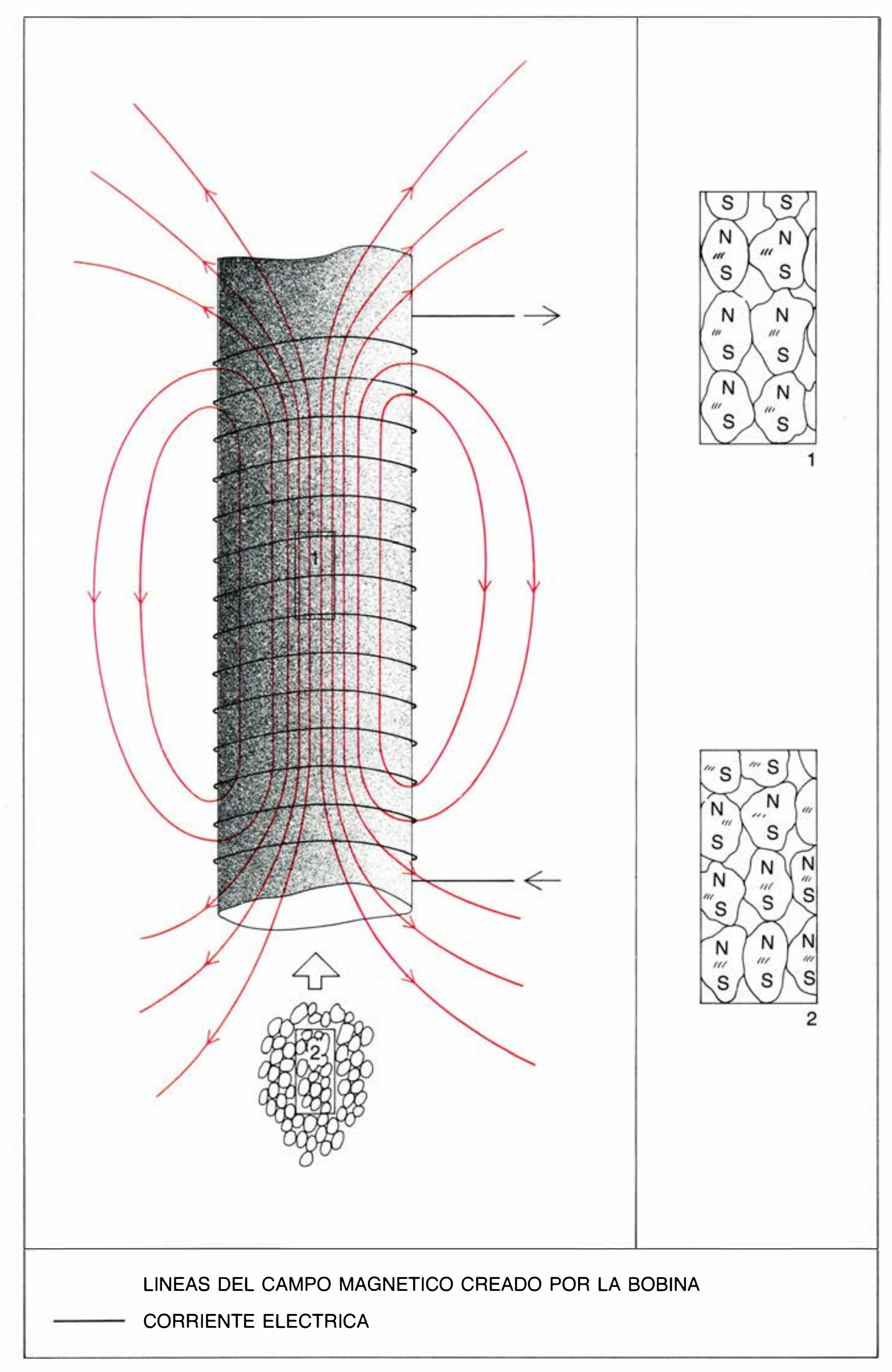
Si colocamos una bobina alrededor de un tubo hueco, el campo magnético creado por la misma en cualquier región del tubo magnetizará cualquier material ferromagnético existente en él, magnetización que dependerá de la proximidad de dicho material a la bobina. Esta magnetización, si es suficientemente intensa, provocará la aglomeración de las partículas. Por otro lado, y debido a esa misma magnetización, si hay partículas ferromagnéticas en el tubo fuera de la bobina y si el campo magnético producido por la misma es suficientemente elevado, tenderán a irse hacia el interior de la bobina, que es la zona donde el campo magnético es mayor. (La intensidad del campo magnético depende de la configuración geométrica de la bobina, el número de espiras y la intensidad de corriente eléctrica que atraviesan ésta.)

Las técnicas de aplicación de campos magnéticos al control de partículas sólidas pueden extenderse al níquel, la magnetita, el hierro y el cobalto hasta que alcancen su temperatura de Curie; esto es, hasta los 358 grados Celsius, 585 grados C, 770 grados C. y 1120 grados C., respectivamente. Cuando estos materiales ferromagnéticos alcanzan la temperatura de Curie, pierden su carácter ferromagnético para transformarse en paramagnéticos. Estos últimos se caracterizan por su debilidad de respuesta ante la acción de los campos magnéticos; necesitarían, pues, unas intensidades de corriente eléctrica prohibitivas desde el punto de vista económico.

Además de materiales ferromagnéticos y paramagnéticos, existen los diamagnéticos, que se distinguen por su tendencia a desplazarse desde zonas de campos magnéticos elevados hasta zonas de campos magnéticos bajos cuando se colocan en campos magnéticos no uniformes. De todos modos, su respuesta ante los campos magnéticos es muy pequeña.

Los materiales paramagnéticos y diamagnéticos son susceptibles de las técnicas descritas aquí. Hay varias formas. En primer lugar, podríamos formar agregados de partículas ferromagnéticas y partículas diamagnéticas o paramagnéticas. Ello supondría la existencia de una unión entre dichos tipos de partículas, química o física, de tal intensidad que no se destruyera en las aceleraciones bruscas y en los choques producidos sobre todo en el TMS. Cabe una segunda posibilidad: formar encapsulados de material magnetizable. En este caso, se prepara el material diamagnético o paramagnético rodeando un núcleo central de material ferromagnético.

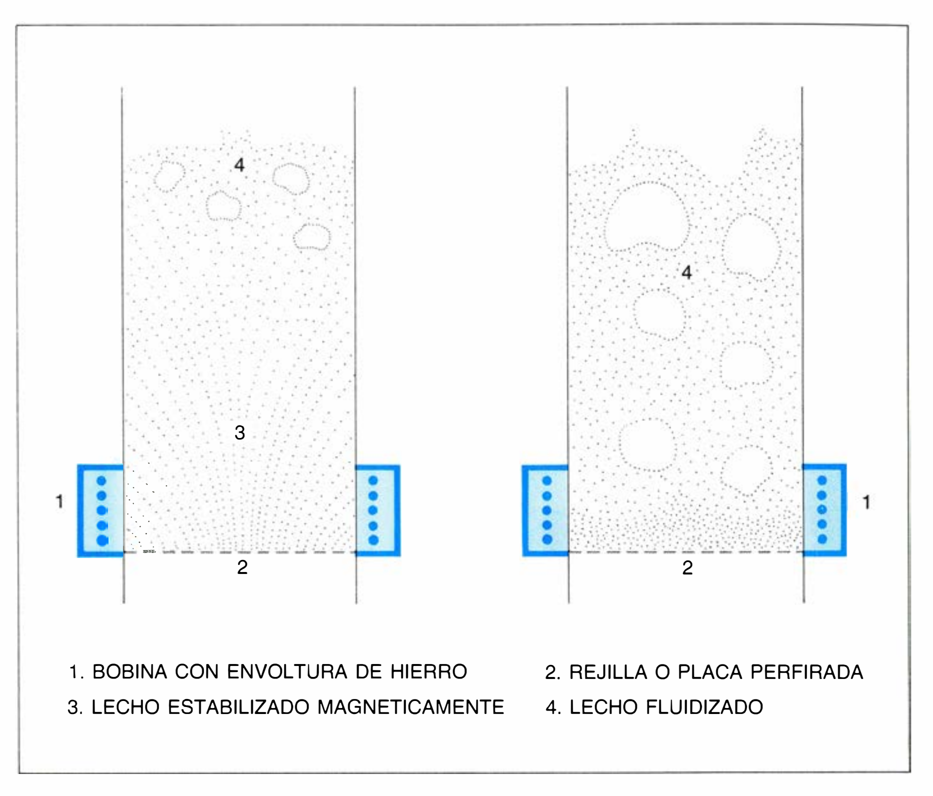
De esos dos modos, por formación de agregados magnéticos o por creación de encapsulados, podemos provocar que los materiales no ferromagnéticos respondan a la acción de cam-



2. FORMACION DE DIPOLOS MAGNETICOS en las partículas ferromagnéticas situadas en el interior o alrededor de una bobina; consta ésta de un conductor por el cual pasa una corriente eléctrica de suficiente intensidad que induce la mencionada formación. Las partículas alojadas dentro de la bobina se unen entre sí al interaccionar mutuamente los dipolos magnéticos, creando cadenas que siguen las líneas del campo magnético. Si el campo adquiere suficiente intesidad, la unión interparticular hará que las partículas se aglomeren y no caigan a través de una rejilla o placa perforada colocada dentro del tubo de la bobina. Al cortar la corriente eléctrica, el campo magnético cesa y las partículas magnetizables caen por la rejilla.

pos magnéticos, puesto que están asociados a una fracción de material ferromagnético. Ambos modos precisan también, sin embargo, una cuidadosa atención para no alterar las propiedades del material diamagnético o paramagnético.

Hay un tercer modo. Consiste en rodear con una malla cerrada de alambre las partículas diamagnéticas o paramagnéticas de interés, junto con la fracción deseada de partículas ferromagnéticas, de suerte que no se puedan salir de la misma. También se puede emplear malla de material ferromagnético, en vez de introducir partículas ferromagnéticas. La malla podría adquirir forma cilíndrica, esférica o cualquier otra deseada. Su tamaño podría oscilar también a voluntad, de acuerdo con su aplicación. Como existe una fracción de material que posee naturaleza ferromagnética, la malla responde a la acción de los campos magnéticos, respuesta cuya intensidad depende del porcentaje de dicho material ferromagnético en la misma. La malla, además, permite el tránsito sin dificultad de los gases; lo que significa que éstos pueden difundirse con libertad hacia las partículas catalíticas situadas en su interior y salir hacia fuera los productos gaseosos. Este tipo de tratamiento no precisa ningún proceso físico o químico que perturbe las propiedades del material no magnético; basta con un dispositivo mecánico que encerrase las partículas ferromagnéticas y no ferromagnéticas en el interior de las mallas. El ámbito de aplicación sería, por fin, general: valdría para toda clase de partículas y el coste del desarrollo sería muy bajo.



3. VALVULA MAGNETICA DE SOLIDOS. Así se llama el mecanismo ideado para aglomeración de las partículas ferromagnéticas situadas en las proximidades de la rejilla y su posterior liberación a través de la misma. Alrededor de la bobina se coloca una envoltura de hierro para evitar la dispersión de las líneas de campo magnético y reducir su consumo energético. Encima de la rejilla se encuentra una zona de partículas formada por un lecho estabilizado magnéticamente y otra constituida por un lecho fluidizado ordinario.

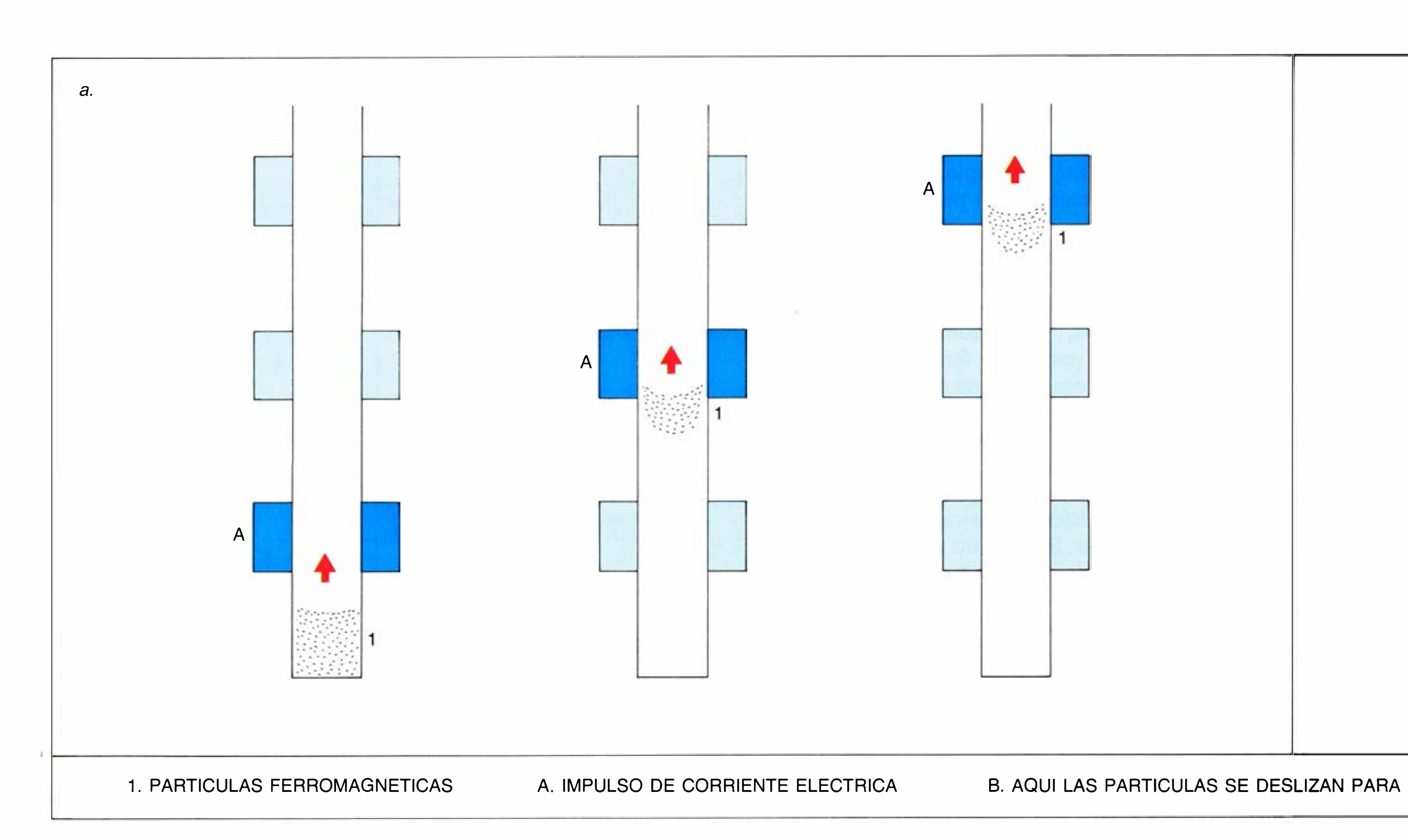
En 1960, el investigador soviético M. V. Fillipov, de la Academia de Ciencias de Latvian, aplicó un campo magnético a un lecho fluidizado de partículas magnéticas. (En el lecho fluidizado las partículas se mueven libremente al paso de un gas, y parte de éste pasa como burbujas.) Para ello colocó una bobina o solenoide alrededor de un lecho empaquetado de partículas o lecho fijo a través del cual fluye una corriente de gases hacia arriba. Al provocar el paso de una corriente eléctrica, las partículas se encuentran en un campo vertical uniforme. Cuando la velocidad de flujo del gas supera la velocidad mínima de fluidización —que es aquella a la cual la fuerza hacia arriba producida por el gas es la misma que el peso de las partículas—, el lecho se expande ligeramente y las partículas se alinean formando cadenas verticales; esta estructura parece inmóvil a la vista, lo que se explica porque el campo es uniforme y las cadenas no se hallan sometidas, en consecuencia, a fuerzas que las hagan ir de una sitio para otro. Este tipo de contactor gas-sólido que utiliza un campo magnético de fondo para estabilizar un lecho fluidizado se denominó lecho estabilizado magnéticamente (LEM).

Con el LEM se pueden poner en contacto un gas y partículas sólidas de un lecho fijo expandido, dotado por consiguiente de elevada porosidad; contacto que puede establecerse a velocidades del gas que provocarían la formación de un lecho fluidizado, si no fuera por la acción estabilizadora del campo magnético.

En 1981, Octave Levenspiel y Thomas Fitzgerald, de la Universidad estatal de Oregon, introdujeron la válvula magnética de sólidos (VMS), que consta de un medio creador de un campo magnético —bobina o conductor eléctrico— y una rejilla para el control de bajada de sólidos por un conducto. La VMS se funda en la propiedad de aglomeración que caracteriza a los sólidos ferromagnéticos sometidos a un campo magnético de suficiente intensidad y en la separación de las partículas al cesar la intervención de dicho campo. En uno de los diseños de dicha válvula, se coloca una rejilla o placa perforada que dejen paso libre a la circulación de las partículas en ausencia de campo magnético. Una bobina rodea el tubo donde se encuentra la rejilla. Al pasar una corriente eléctrica de intensidad suficiente, las partículas se aglomeran y no atraviesan la rejilla. Sí pasan por ella cuando cesa la corriente. Podemos regular los tiempos de activación de la corriente para conseguir los caudales deseados de partículas sólidas.

Cuando se inyecta un gas en dirección ascendente se forma una zona en LEM, en cuya parte superior, donde el campo no adquiere especial intensidad, existe un lecho fluidizado. Cortemos la corriente eléctrica y las partículas atravesarán el contactor permitiendo el control de la caída de las mismas mediante la regulación de los intervalos de corte.

Se dispone, pues, de la combinación de una placa distribuidora de flujo y de una válvula para el control de las partículas. Entre las grandes ventajas que presenta la VMS citaremos la ausencia de partes móviles, con lo cual se elimina la necesidad de mantenimiento y la posibilidad de fallos, y el tiempo de respuesta del sistema a los cambios de control, que es muy rápido. Asímismo, se pueden tratar mezclas de partículas



4. TRANSPORTADOR MAGNETICO DE SOLIDOS diseñado por el autor para mover partículas en sentido horizontal. Supone una ventaja sobre el sistema neumático habitual: no requiere corrientes de gases auxiliares, pudiendo formar parte de sistemas completamente cerrados de recirculación de partículas; gracias a ello se elimina la necesidad de compresores, soplantes, motores y demás para hacer circular los gases. El movimiento tiene una componente

paramagnéticas y diamagnéticas, siempre que exista cierta fracción de partículas ferromagnéticas.

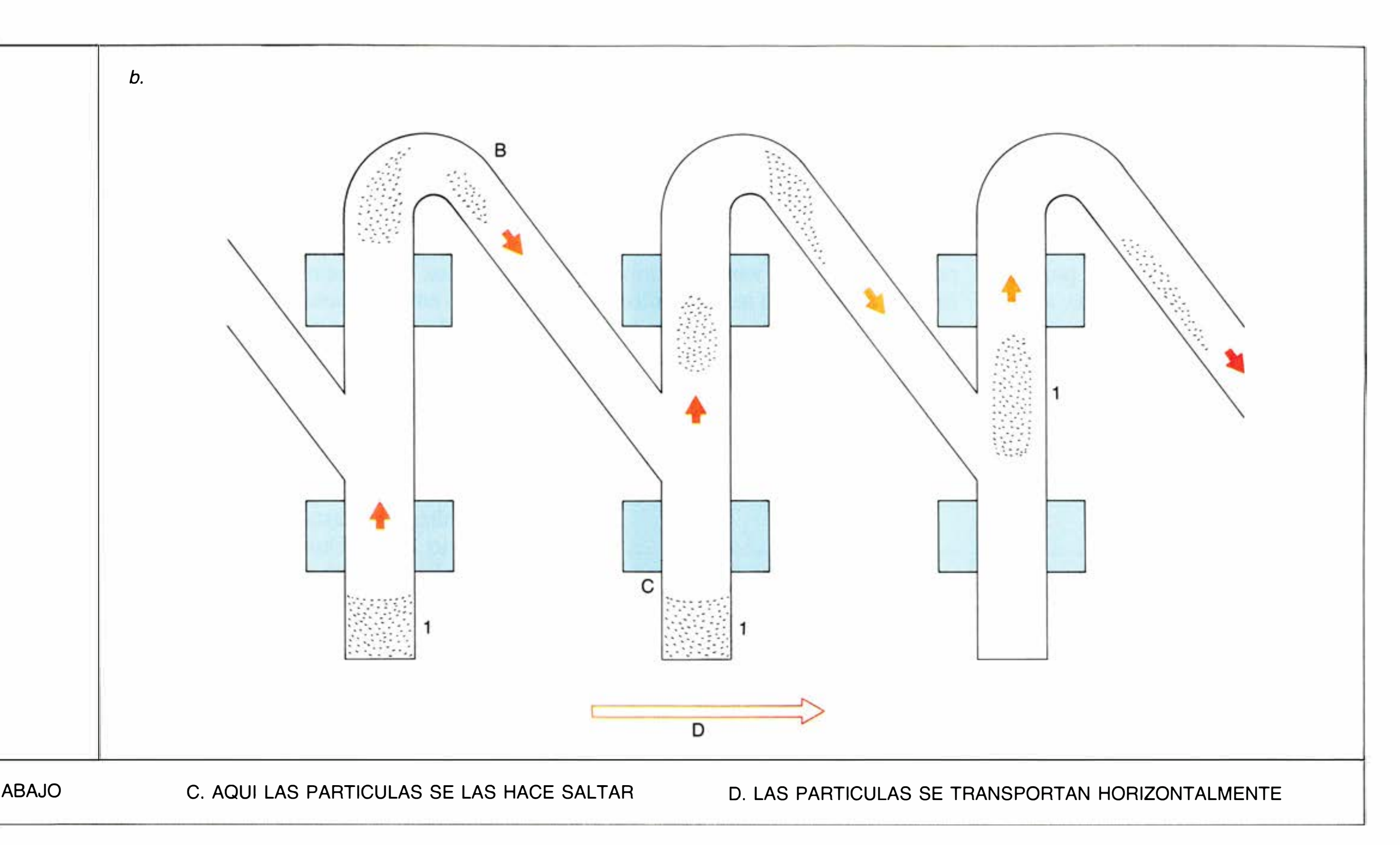
En 1983, en la Universidad de Salamanca, diseñamos una nueva técnica, capaz de utilizar campos magnéticos variables para mover partículas sólidas en sentido vertical sin recurrir a los medios neumáticos. Lo llamamos transportador magnético de sólidos (TMS). El primer diseño se fundaba en el empleo de solenoides con núcleo ferromagnético, situados a ambos lados de un conducto rectangular; distribuíamos los solenoides alternativamente con la esperanza de que, merced a su conexión secuencial en el tiempo, las partículas ferromagnéticas irían saltando de uno a otro en una trayectoria en zigzag. El experimento fracasó por la presumible acción perturbadora de los núcleos ferromagnéticos de los solenoides sobre las líneas de campo magnético.

Más tarde, creamos otro diseño que funcionó satisfactoriamente. Consta de una serie de bobinas colocadas de forma espaciada alrededor del tubo por donde las partículas han de ascender verticalmente. Así, en el supuesto de que las partículas magnéticas se hallen en un comienzo junto a la primera bobina, se provoca el paso de un impulso de corriente continua, de varias decenas de ampère y pocos volt, durante cierto tiempo; el intervalo en cuestión dependerá de la intensidad de la corriente eléctrica, de la naturaleza de las partículas y del diámetro del tubo. Tiempos típicos de los impulsos son de ochenta a cien milisegundos.

Las partículas, que reciben una intensa aceleración, se encaminan hacia el interior de la bobina; mas, a medida que se van aproximando al centro, cesa el impulso de la corriente, y las partículas pueden proseguir su vía ascensional hacia la segunda bobina. Al no existir ya fuerza de atracción magnética en la región situada entre la primera y la segunda bobina, la gravedad frena las partículas. Pero eso no impide que se acerquen hasta las inmediaciones de la segunda bobina; se provoca entonces otro impulso de corriente eléctrica que las acelera de nuevo, repitiéndose el proceso. Así puede desarrollarse el transporte a lo largo de un tubo en sentido vertical cuando se dispone del número adecuado de bobinas.

El sistema TMS posibilita, a su vez, el transporte horizontal de partículas. Con ese propósito, se dispone que las partículas asciendan y circulen luego a través de un codo, deslizándose por un tubo inclinado en cierto ángulo. Cuando las partículas llegan al codo, otro TMS se encarga de elevarlas de nuevo, y así sucesivamente. Podría pensarse en la incorporación de bobinas para el transporte horizontal y directo de partículas, pero entonces habría que contrarrestar la fuerza de la gravedad, amén de que las partículas, con tales bobinas, se dispersan mucho.

Los primeros trabajos experimentales de J. A. Briz y el autor sobre el TMS se efectuaron con tubos de hasta 30 milímetros de diámetro. Estudiamos las influencias, en el sistema, del diámetro del tubo, de la naturaleza de las partículas, de la distancia entre las bobinas y de la duración de los impulsos eléctricos. Gerardo Martín y el autor ampliaron el ensayo empleando bobinas con diámetro de hilo bastante mayor, con lo que se gastaba casi la misma energía y, sin embargo, se obtenía un número mayor de espiras. Utilizamos tubos de hasta 50 milímetros, con lo que la cantidad transportada por impulso llegó hasta 160 gramos de acero inoxidable. (Por mor de comparación, en los primeros ensayos sólo se podían transportar dos gramos.)



ascencional y otra descendente a lo largo de un tubo inclinado; para el movimiento ascensional se utiliza un transportador que toma las partículas situadas en la parte inferior del tubo y las lleva a la parte superior, donde existe un codo con formas redondeadas que comunica con un tubo inclinado descendente conectado con otro transportador. La morfología del codo resulta decisiva para el curso de las partículas. Los impulsos magnéticos deben estar sincronizados.

A la vista de esos resultados determinamos abordar con mayor profundidad las variables que condicionaban los saltos de las partículas ferromagnéticas, en particular el diseño de la bobina. Cualquier proyecto de ésta ha de preguntarse por su contribución al mayor rendimiento de los saltos de las partículas. Como es obvio, la eficacia ideal se da cuando todas las partículas saltan a la vez y ascienden a la misma velocidad. Dicho de otra manera, que la dispersión de las partículas es mínima. Había, pues, que considerar distintas posibilidades de bobina para un determinado diámetro de tubo.

Partimos de la situación en que las partículas se encuentran en reposo, dentro del tubo cerca de la primera bobina, y ocupan cierto volumen. La primera condición que se impone es la relativa a la potencia de la fuente de alimentación en corriente continua. Decidida dicha potencia, se estudian las diferentes posibilidades de bobinas, esto es, las varias combinaciones de voltaje, intensidad de corriente, diámetro del conductor que forma la bobina, número de espiras por capa y número de capas de espiras. En definitiva, se trata de ajustar la bobina a los dos criterios de optimización siguientes: primero, valor máximo del producto del campo magnético por el gradiente del mismo promediado en toda la zona ocupada por las partículas, y, segundo, valor mínimo de la diferencia de los ángulos que forman la dirección de las fuerzas ejercidas por la bobina en relación con la vertical en las zonas cercanas al eje del tubo del TMS y las zonas cercanas a la pared del tubo. Por lo que concierne al primer criterio, la fuerza ejercida sobre las partículas es directamente proporcional al producto mencionado; por lo que se refiere al segundo, cuanto menor sea la diferencia de ángulos, menor será la tendencia que muestren las partículas a diversificar sus caminos.

Una vez escogidos el tipo óptimo de bobina para un determinado diámetro del tubo del TMS y la fuente de alimentación, se simula el funcionamiento del TMS para estimar el tiempo de los impulsos, así como la cantidad de partículas por impulsos que puede transportar. Comprobamos con ello que el tipo de bobina y la fuente de alimentación son los adecuados.

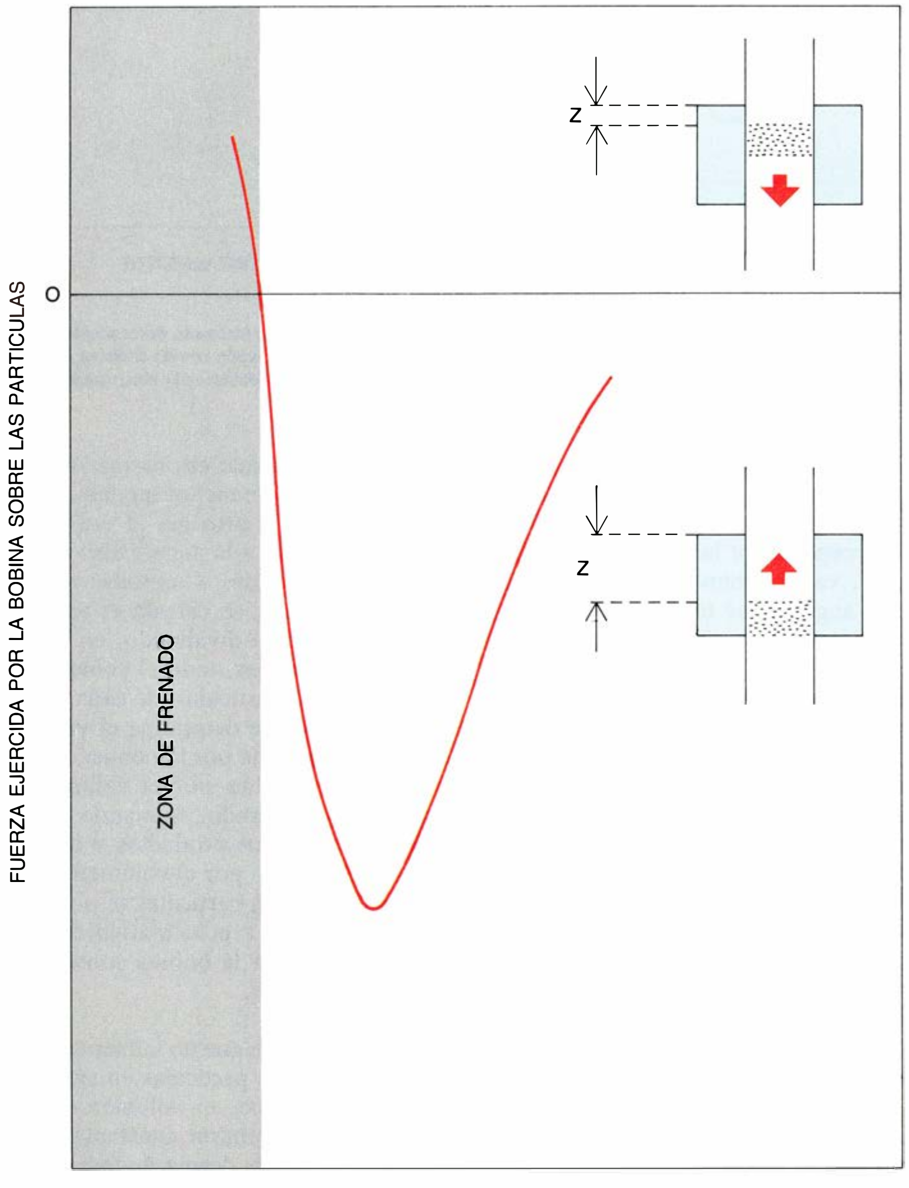
A tal fin, se empieza por definir un valor medio de la fuerza que ejerce la bobina sobre las partículas. Fuerza que varía bastante con la distancia de las partículas a la bobina; esa fuerza varía también, aunque en menor medida, de la zona interior y próxima al eje del tubo con respecto a la zona exterior y próxima a las paredes. Con todos esos datos por delante, se calcula el valor medio de la fuerza dividiendo, en volúmenes elementales, todo el volumen ocupado por las partículas; de cada volumen elemental se determina el valor de la fuerza ejercida por la bobina, valor que se multiplica por el volumen elemental considerado. Sumando los valores de todos los productos y dividiendo el resultado por el volumen total ocupado por las partículas se obtiene un valor medio representativo de la fuerza ejercida por la bobina sobre el grupo de partículas.

En la hipótesis de que no hubiera dispersión de las partículas en su ascensión por el tubo, el volumen que ocuparan se mantendría constante, si bien su posición iría desplazándose hacia arriba. Si las partículas arrancaran de determinada distancia de la bobina y la fuerza ejercida por ésta fuera lo bastante intensa para provocar el salto de las partículas, la fuerza iría aumentando a medida que las partículas se fueran acercando a la bobina. Llegadas

al interior de ésta, se produce una inflexión en el valor de la fuerza frente a la distancia. En efecto, a partir de entonces, la fuerza disminuye rápidamente a medida que las partículas se desplazan dentro de la bobina hasta que alcanzan la zona de frenado, donde, de nuevo, la fuerza torna a invertirse: cambia su sentido ascendente por otro descendente, que se suma a la acción de la gravedad. Antes de alcanzar la posición en que se produce la inversión de la fuerza, hemos de cortar la corriente eléctrica que pase por la bobina y permitiremos así que las partículas prosigan su camino ascendente, no se frenen y no se adhieran al tubo.

Al representar gráficamente la aceleración a la que se ven sometidas las partículas en su ascensión por el tubo de TMS mientras dura el impulso de la corriente eléctrica, descubrimos un pico o máximo de la misma. En un comienzo, las partículas, alejadas de la bobina, apenas notan la aceleración, pero conforme se van aproximando a ésta, la aceleración se intensifica hasta que alcanza su punto máximo cerca de la zona de frenado; a partir de esa cota, cae bruscamente.

Otra magnitud interesante es la altura del salto de las partículas. Cuando se calcula hasta dónde llegan en el mismo, en función del tiempo de duración del impulso eléctrico, se observa que la altura máxima de los saltos se produce cuando el impulso dura hasta que las partículas se aproximan a la zona de frenado, sin entrar todavía en ella. De ahí se desprende la importancia de la sincronización de los impulsos eléctricos para que no se produzcan frenados y que los saltos sean grandes. (En el caso de emplear bobinas de mayor potencia, la fuerza producida sería también mayor, las partículas llegarían antes a la zona de frenado y el tiempo de impulso necesario sería a su vez menor.)



5. EVOLUCION DE LA FUERZA EJERCIDA POR LA BOBINA sobre el volumen total de partículas ferromagnéticas y su representación en función de la distancia entre las partículas y el extremo superior de la bobina. Inicialmente, cuando las partículas están en las proximidades de la bobina aunque relativamente alejadas, la fuerza media de atracción al interior de la bobina es pequeña; su valor es negativo porque se trata de una fuerza de atracción. A medida que van ascendiendo y acercándose a la bobina, aumenta la fuerza de atracción hasta que se llega a un valor máximo, a partir del cual la fuerza disminuye rápidamente. Aparece entonces una fuerza hacia abajo que va aumentando. Las partículas entran en la zona de frenado.

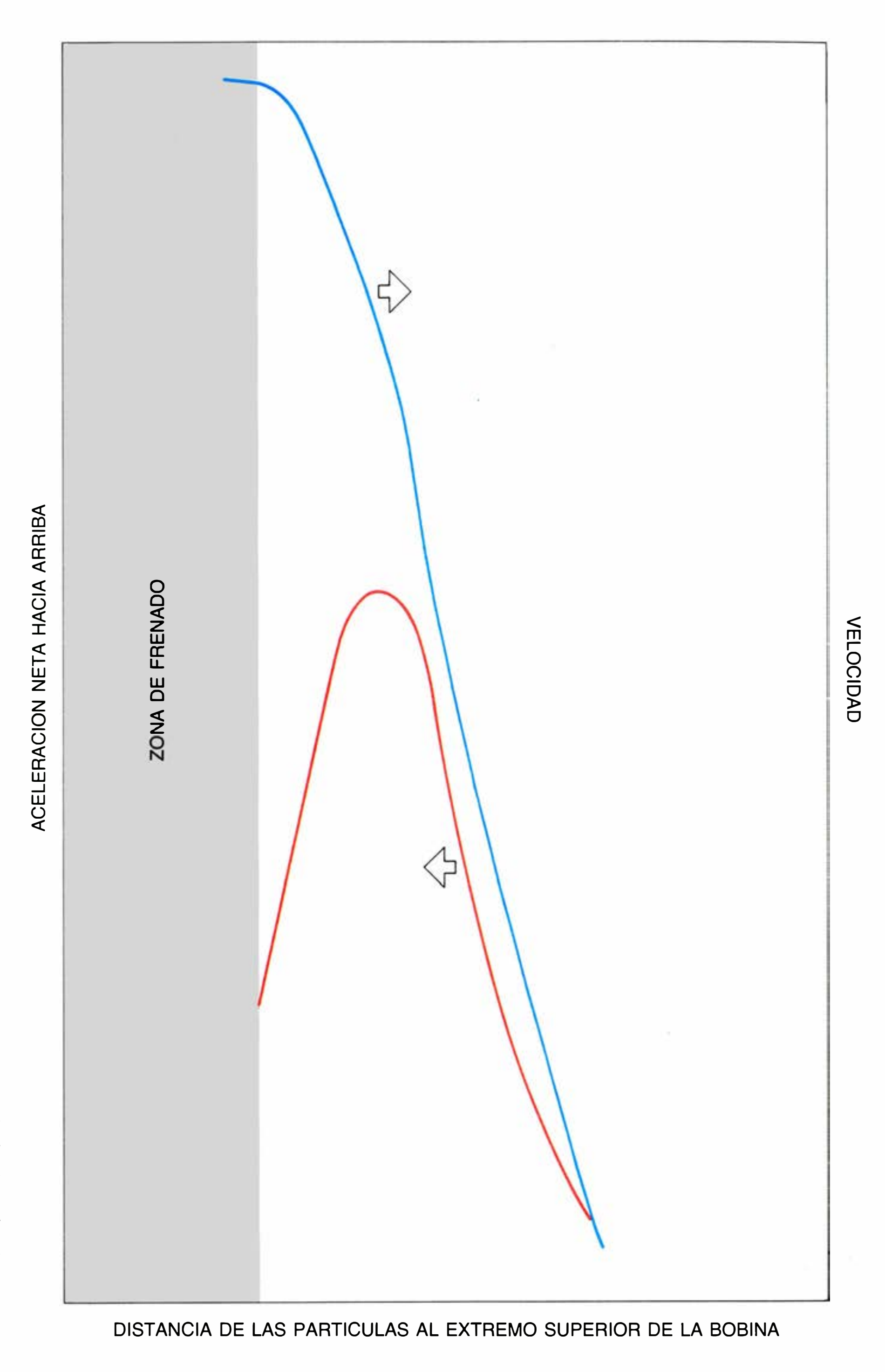
En los ensayos que hemos venido realizando hasta ahora en mi laboratorio de la Universidad de Salamanca, la sincronización de los saltos se ha efectuado de forma empírica, variando los tiempos de impulso hasta que los saltos nos parecieron satisfactorios. En el caso de un tubo de un TMS de 50 milímetros de diámetro se consiguieron alturas superiores a los 600 milímetros, aunque se producía también una gran dispersión de partículas. (Partículas que ocupaban en un comienzo una zona de altura de 30 milímetros en el tubo, tras el primer salto se dispersaban en un radio de más de 200 milímetros, cuando se encontraban ya en la región de los 420 a los 620 milímetros de distancia de la base. Esto sucedía con tiempos de impulso de 80 milisegundos y una masa de partículas de hierro de 125 gramos. Sin embargo, cuando el tiempo del impulso crecía hasta los 103 milisegundos en las mismas condiciones, la dispersión de las partículas se reducía a la mitad.) El problema de la dispersión requiere especial cuidado porque, si las partículas llegan desparramadas a la segunda bobina, ésta sólo podrá impulsar la parte superior de la zona ocupada por las partículas y las otras caerán capturadas en la propia bobina sin capacidad de ascender tubo arriba; lo que, a su vez, complicaría el mecanismo de control.

En 1986, G. Martín y el autor consiguieron, sin embargo, recircular continuamente partículas a través de las nueva bobinas de un TMS cuyo tubo era de 20 milímetros de diámetro. ¿Cómo solucionar el problema de la dispersión? El método que ideamos se apoya en el empleo de bobinas impulsoras y bobinas detectoras. Las segundas producen una pequeña fuerza electromotriz, en cumplimiento del principio de inducción electromagnética de Faraday, cuando las partículas ferromagnéticas pasan por su interior. Las bobinas detectoras revelan la posición

en que se hallan, en cada momento, las partículas que transportan las bobinas impulsoras. Con esa información, un ordenador, dotado de un programa de control, podría establecer la conexión o desconexión en el momento oportuno de cada bobina impulsora del sistema.

La combinación de la técnica de subir las partículas sólidas por medio del transportador magnético de sólidos y la técnica de controlar su caída y ponerlas en contacto con un gas en la válvula magnética de sólidos constituye el recirculador magnético de sólidos (RMS). Son múltiples sus aplicaciones potenciales. Aquí nos limitaremos a exponer dos: el aprovechamiento del calor de un gas, verbigracia el procedente de una cámara de combustión para precalentar el aire de combustión, y la recirculación de partículas catalíticas que se desactivan rápidamente.

En el caso de un recuperador de calor de los gases de combustión se dispondría como mínimo de dos contactores con válvulas magnéticas para sólidos. El contactor gas-sólido, situado en la parte superior del RMS, serviría para recuperar el calor de los gases de combustión. Para ello, por la parte superior del contactor entrarían partículas ferromagnéticas frías que constituirían un lecho magnéticamente estabilizado encima mismo de la válvula magnética; podría tener, igualmente, un lecho fijo o un lecho fluidizado de partículas, según fuere la velocidad de los gases a través del lecho, la disposición del medio creador del campo y la intensidad de este último. Al poderse utilizar partículas pequeñas, la relación entre el área superficial de las partículas y la masa de las misma da un valor alto y, en consecuencia, la velocidad de transferencia de calor es elevada. En otras palabras, con un volumen limitado podríamos obtener una transferencia de calor considerable.

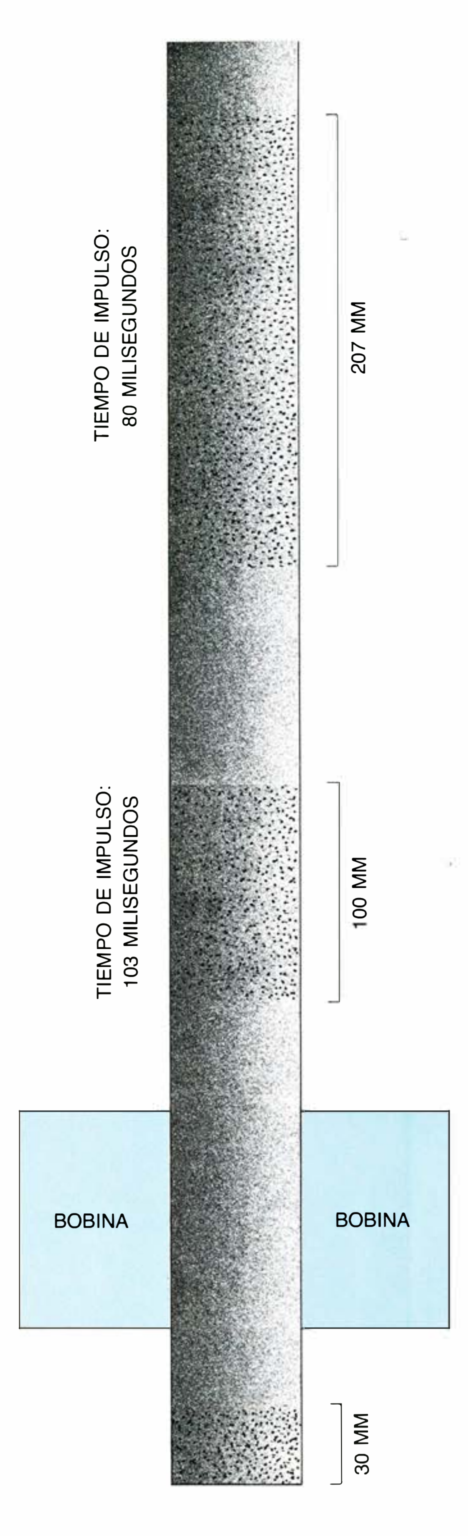


6. DESPUES DEL INSTANTE INICIAL en que las partículas ferromagnéticas han saltado, al aproximarse cada vez más a la bobina la fuerza media de atracción ejercida por la bobina aumenta. La aceleración neta hacia arriba —es decir, la aceleración producida por la bobina menos la aceleración de la gravedad— aumenta hasta alcanzar un máximo. Luego, disminuye rápidamente. Ese máximo se alcanza en la región del interior de la bobina. Una vez que las partículas comienzan a introducirse allí, las que van en cabeza se convierten en obstáculo para las siguientes, se crea una fuerza de frenado que disminuye la fuerza neta hacia arriba.

A medida que las partículas que entran en contacto con el gas caliente se fueran calentando, la válvula magnética de sólidos facilitaría su paso al otro contactor que está debajo. Ello permitiría siempre la renovación del frente de partículas sometidas a calentamiento. Por otro lado, las partículas que cayeran hacia el contactor inferior, dotado de otra válvula magnética para sólidos, se irían poniendo en contacto con el aire que se fuera a precalentar, saliendo éste por un conducto distinto del de caída de las partículas. El aire podría pasar entonces a la cámara de combustión.

Llegadas las partículas a la rejilla de la válvula magnética de sólidos, se enfrían; caen entonces por las mismas y arriban hasta la parte inferior del TMS; allí son recogidas por el propio transportador y transferidas a la parte superior del primer contactor, que es el recuperador de calor de los gases de combustión. Este sistema de recuperación del calor de los gases, si aplicamos campos magnéticos que impidan el arrastre de las partículas por los gases, constituyen un sistema cerrado de transferencia de calor que no precisa de partes móviles para hacer recircular las partículas.

Veamos la segunda aplicación: empleo del reactor magnético de sólidos



7. SALTOS PRODUCIDOS en un transportador magnético de sólidos de 50 milímetros de diámetro cuando el impulso eléctrico es de 48 ampère. En la bobina inicial se produce una dispersión de las partículas, que depende de la distancia de la bobina a las partículas; en este caso se ha cifrado en 30 milímetros. La dispersión del volumen inicial y la altura que alcanzan tienen que ver con los tiempos de duración del impulso. Conforme esta última se hace más corta, las partículas se elevan más, aunque se hallan también más dispersas. Para un tiempo de 80 milisegundos, se ha saltado una distancia de 207 milímetros, estando situado el extremo inferior de las partículas a unos 420 milímetros. Un diseño optimizado de la bobina, junto con la sincronización más perfecta de los impulsos magnéticos, posibilitará que la dispersión disminuya mucho. Los experimentos realizados indican que este problema se presenta, de forma casi exclusiva, en las primeras bobinas.

para establecer un circuito cerrado de catalizador sin la necesidad de medios neumáticos de transporte y con las mismas ventajas de la aplicación anterior. En este caso, el catalizador se encontraría con el gas en el contactor superior. En efecto, el gas saldría por dicha parte superior una vez hubiera reaccionado. El lecho, donde caería el catalizador, podría ser, de nuevo, magnéticamente estabilizado, fluidizado o fijo. La parte del catalizador situada encima mismo de la VMS, al estar en contacto con el gas reaccionante que apenas se ha convertido en productos, se desactivaría antes (es decir, deja de ser catalizador para convertirse en sólido inerte). Una vez desactivado el catalizador, se le franquea el paso hacia el contactor de regeneración, y allí sufriría un nuevo tratamiento que lo convertiría en catalizador activo. Este pasaría, a través de una válvula magnética, a la primera bobina del transportador magnético de sólidos, que lo conduciría nuevamente al reactor.

Las investigaciones llevadas a cabo y otras en curso en nuestro laboratorio han puesto de manifiesto la posibilidad de controlar el movimiento de los sólidos por medio de campos magnéticos. En particular, el desarrollo del transportador magnético de sólidos, que permite el ascenso de partículas sólidas por medio de campos magnéticos variables, ha abierto el camino del recirculador magnético de sólidos, efectuándose el control de la caída de las partículas por medio de la válvula magnética de sólidos.

Se logra con ello la creación de circuitos cerrados de partículas sólidas, de previsible utilidad futura en el campo de la transferencia de calor y en la activación-desactivación de los catalizadores. Uno de los problemas que queda por resolver, y en el que estamos trabajando, es el relativo al desarrollo del distema de control de recirculación, así como de su optimización.

BIBLIOGRAFIA COMPLEMENTARIA

THE USES OF MAGNETIC FIELD IN THE PROCESSING OF SOLIDS. E. Jaraíz, O. Levenspiel y T. J. Fitzgerald en *Chemical Engineering Science*, vol. 38, págs. 107-114; 1983.

TRANSPORTE MAGNÉTICO DE SÓLIDOS. E. Jaraíz, J. A. Briz y M. Jaraíz en *Ingeniería Química*, n.º 203, págs. 103-111; febrero de 1986.

DESARROLLO DE UN RECIRCULADOR MAGNÉTICO DE SÓLIDOS. G. Martín y E. Jaraíz en *Química e Industria*, vol. 33, n.º 12, págs. 1055-1062; diciembre de 1987.

Activadores génicos

Los principios de los mecanismos de activación y desactivación de genes en células bacterianas se conocen ya con bastante detalle. ¿Podemos aplicar esos principios a la regulación génica en organismos superiores?

Mark Ptashne

Cada célula humana contiene unos 100.000 genes, pero no todos están siempre activos. Muchos se expresan selectivamente, durante ciertas etapas del desarrollo, por ejemplo, o en respuesta a señales ambientales. ¿Cómo se activan y desactivan los genes? Durante los últimos veinte años, el autor y sus colegas han estudiado, en la Universidad de Harvard y en otros laboratorios, los mecanismos responsables de esos cambios en un virus bacteriano, el bacteriófago lambda. Las ideas desarrolladas con ese sencillo modelo permiten avanzar con rapidez, lo vemos ahora, en el análisis de la regulación génica de otras formas de vida más complejas, como las células humanas.

Para entender el problema es preciso recordar que un gen típico, un segmento de ADN, determina una proteína. La secuencia específica de pares de bases en ese segmento de ADN se corresponde con la secuencia específica de aminoácidos que forman la proteína. Descifrar un gen requiere, primero, su transcripción en una molécula de ARN mensajero (ARNm) y, después, la traducción de ésta en una proteína.

La expresión génica puede controlarse por proteínas reguladoras que se unen a sitios específicos del ADN. Estas proteínas se denominan activadores o represores, según incrementen o disminuyan la tasa de transcripción; algunos reguladores realizan las dos funciones. ¿Cómo reconocen las proteínas reguladoras esos sitios específicos del ADN y activan o desactivan los genes que controlan?

Las polimerasas de ARN, enzimas encargadas de transcribir los genes en ARN, intervienen en este proceso. En el caso de los genes que nos interesan aquí, una polimerasa de ARN se une al ADN en un sitio cercano al comienzo de un gen, y, luego, se mueve a lo largo del mismo, transcribiendo la secuencia de ADN en una molécula de ARNm. Sabido esto, la pregunta se torna más precisa: ¿qué tipo de interacción se produce entre un activador y la polimerasa de ARN para que ésta transcriba un gen? ¿Cómo impide un represor que la enzima transcriba?

Los estudios sobre el fago lambda nos enseñaron mucho sobre interacciones entre reguladores, ADN y polimerasa de ARN en los procesos de regulación génica de este organismo sencillo. Teníamos, no obstante, bastantes razones para sospechar que nuestros descubrimientos no serían aplicables a organismos superiores. En primer lugar, las polimerasas de ARN de los organismos inferiores diferían bastante, a primera vista, de las propias de los organismos superiores. También, mientras los reguladores que estudiamos en lambda se unían a sitios muy cercanos a los genes que controlan, en otros casos los sitios a los que se unían los reguladores quedaban alejados del gen cientos o incluso miles de pares bases. Además, en eucariotas (organismos superiores), a diferencia de los procariotas, como las bacterias, el ADN está encerrado en un núcleo, y se enrolla alrededor de unas proteínas llamadas histonas, como el hilo en la bobina. ¿Existen mecanismos fundamentalmente diferentes en estos casos?

Los resultados obtenidos en los dos últimos años sugieren la existencia de unos pocos principios sencillos, comunes a los procesos de regulación génica en situaciones tan distintas como ésas, ya sea en bacterias, levaduras, plantas, moscas de la fruta o el hombre. Expondré estas ideas tal como fueron surgiendo de los estudios realizados con el fago lambda, y mostraré cómo, junto con mis colegas, estamos aplicándolas al estudio de la regulación génica en organismos superiores. En particular, me detendré en el problema de la activación génica.

El fago lambda centró la atención de los científicos hace más de treinta años, cuando éstos se percataron de que el ciclo de vida del fago constituía un ejemplo claro de regulación génica. Cuando se inyecta ADN vírico en el interior de la bacteria, sigue uno de estos dos caminos: o se expresan la mayoría de los genes víricos y el virus se replica y lisa a la célula hospedadora, o casi todos los genes se inactivan y el ADN vírico, inerte, se convierte en parte pasiva del cromosoma bacteriano. Aun así, los genes silenciados se activan cuando la bacteria se tropieza con un carcinógeno, la luz ultravioleta por ejemplo.

Una proteína reguladora determinada por el virus desempeña un papel clave en el control de la expresión de los genes víricos. Esta proteína, el represor de lambda, se une a unos sitios específicos del cromosoma vírico y desactiva la mayoría de los genes víricos. Los carcinógenos, destruyendo al represor, inducen la expresión de los genes responsables de la replicación del virus y de la lisis celular (los genes "líticos"). Además de ser un represor de la expresión génica, como indica su nombre, esta proteína *activa* también la transcripción; dicho de otra manera: incrementa notablemente la tasa de transcripción de su propio gen.

El conocimiento de la estructura del represor nos ha ayudado a entender su mecanismo de acción. La unidad fundamental del represor, o monómero, se pliega para generar una suerte de burbujas (dominios), de tamaño parejo. La asociación de dos monómeros origina un dímero, forma en que el represor se une al ADN. Son dos los dímeros que se unen a sitios contiguos del ADN de lambda: O_R1 y O_R2; cerca, hay otro sitio más, el O_R3, pero en nuestra exposición podemos dejarlo de lado.

MARK PTASHNE es profesor de bioquímica y biología molecular en la Universidad de Harvard. Este es su segundo artículo sobre regulación génica que ha escrito para INVESTIGACIÓN Y CIENCIA.

La presencia de ambos dímeros unidos al ADN afecta a la transcripción de dos maneras, una negativa y otra positiva. Primero, los represores impiden que la polimerasa de ARN se una al ADN y copie los genes líticos (que por convención se dice que están a la derecha). Segundo, uno de los represores facilita la unión de la polimerasa de ARN para que transcriba el gen que determina la síntesis del propio represor (situado a la izquierda de donde están unidos los represores).

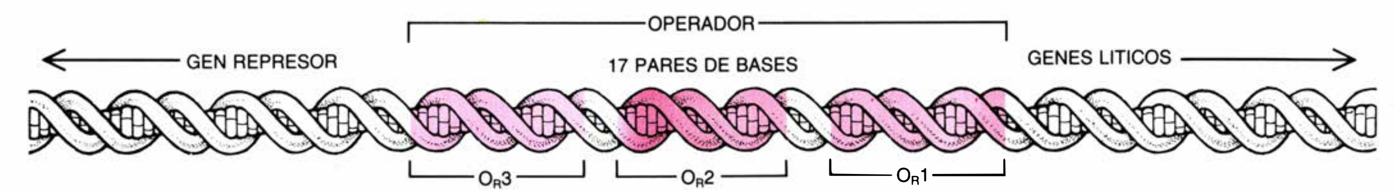
Los sitios a los cuales se une el represor se llaman operadores y, promotores, aquellos a los que se une la polimerasa de ARN. El promotor situado a la derecha del operador de lambda es fuerte y el de la izquierda, débil. Los sitios operadores a los cuales se une el represor se solapan con el promotor fuerte de la derecha, pero no así con el débil de la izquierda, que queda adyacente a los mismos. En ausencia de represor, la polimerasa de ARN se une al promotor fuerte de la derecha y comienza la transcripción. Por el contrario, sólo si es ayudada por el represor, cuando está unido en la zona contigua (en su papel de activador), puede la polimerasa unirse al promotor débil y transcribir el gen del represor. Una pequeña diferencia en las secuencias de los dos promotores explica el hecho de que un promotor atraiga sólo débilmente a la polimerasa, y requiera por tanto la presencia de un activador para cumplir su misión, mientras que el otro funciona bien, siempre que la polimerasa pueda acceder a él.

Hay una sólida razón biológica para justificar la complejidad de este cuadro: los dos dímeros represores se unen simultáneamente y el represor tiene un doble efecto, positivo-negativo. Estos y otros factores permiten que el virus pase, eficazmente, de un modo de crecimiento a otro, dependiendo del ambiente [véase "Un interruptor genético en un virus bacteriano", por Mark Ptashne, Alexander D. Johnson y Carl O. Pabo; INVESTIGACIÓN Y CIENCIA, enero de 1983]. Volvamos ahora a nuestra pregunta original, ya de una forma más precisa: ¿de qué manera se une el represor a sus sitios operadores específicos y activa la transcripción de su propio gen?

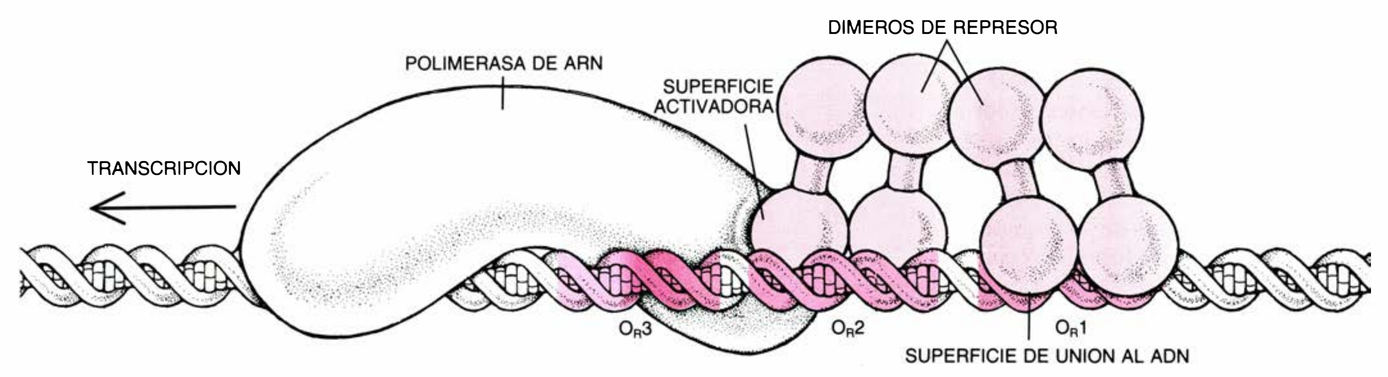
La cristalografía de rayos X y múltiples experimentos bioquímicos han auspiciado una visión pormenorizada de la unión específica del represor de lambda, y otras proteínas reguladoras parecidas, a sus sitios operadores en el ADN. La superficie del monó-



1. LARVA DE MOSCA DE LA FRUTA, teñida de azul en el intestino medio anterior (*arriba*) y en el cuerpo graso (*abajo*), lo que revela la actividad de una enzima implicada en el metabolismo de azúcares. Una proteína reguladora que normalmente se encuentra en levaduras puso en acción al gen responsable de la enzima. El activador de levaduras se halla capacitado para operar en células de plantas, mamíferos e insectos.



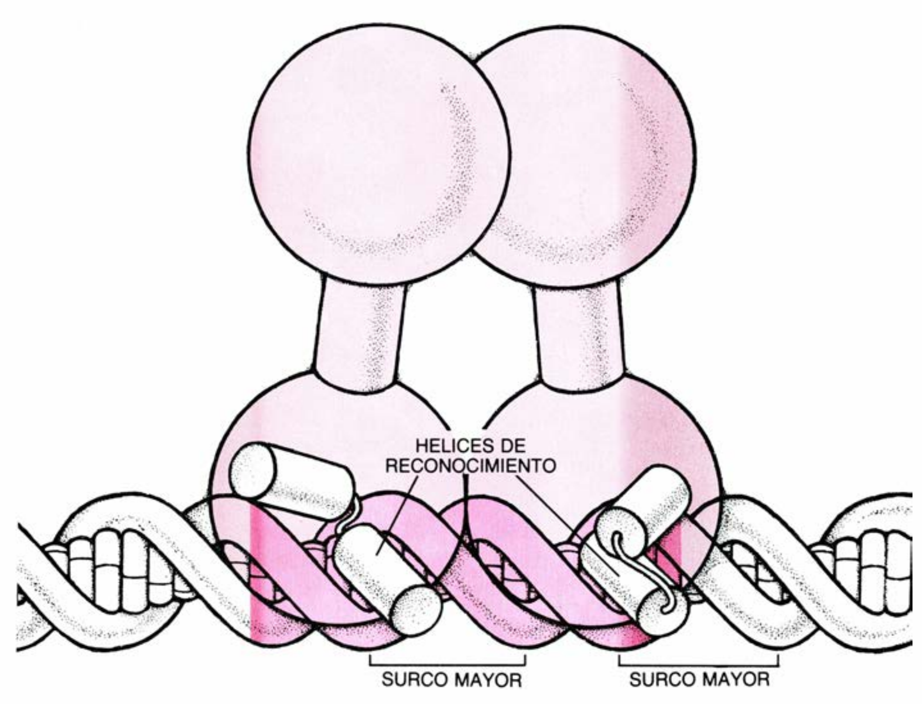
2. OPERADOR DE LAMBDA, con los sitios donde se unen las proteínas reguladoras que controlan los genes del bacteriófago lambda. El gen de una proteína reguladora denominada represor se sitúa a la izquierda del operador; a la derecha del operador están los genes líticos, que intervienen en la replicación de los virus y en la lisis de la bacteria hospedadora. El operador aloja tres sitios de unión, designados O_R1, O_R2 y O_R3, cada uno de ellos de 17 pares de bases de ADN.



3. EL REPRESOR ACTIVA SU PROPIO GEN coadyuvando a que la enzima polimerasa de ARN se una al gen represor. La polimerasa de ARN transcribe genes en ARN, primer paso del proceso de expresión de un gen en proteína. Los "dímeros" de represor se unen a O_R1 y O_R2. (Un dímero es una molécula formada por dos subunidades idénticas.) Se piensa que una parte de la proteína, la "superficie activadora", hace contacto con la enzima. Los dímeros se yuxtaponen también unos a otros, para facilitar su propia unión con el ADN, un fenómeno conocido por cooperatividad. Además, los dímeros, cuando están unidos al ADN, bloquean la transcripción de los genes líticos, e impiden que la polimerasa se una al ADN. Por ello, la proteína recibe el nombre de "represor".

mero-represor que se une al ADN presenta una hélice alfa. Son muchas las proteínas que poseen esas protuberancias, y son muchas también las secuencias de aminoácidos que se pueden plegar adoptando configuraciones en hélice alfa.

La hélice alfa de cada monómero encaja a la perfección en el surco mayor de la doble hélice de ADN. Los grupos químicos situados en la superficie exterior de la hélice alfa forman un patrón característico, que viene determinado por la secuencia de aminoácidos que integran la hélice alfa. Los grupos químicos de los pares de bases expuestos en el surco mayor forman también un patrón, determinado por la secuencia de pares de bases. Sólo cuando ambos patrones casan se puede unir el represor [*véase la figura 4*].



4. SUPERFICIE DE UNION AL ADN DEL REPRESOR, que interacciona con el ADN a través de ciertas estructuras sobresalientes, las hélices alfa, formadas por cadenas helicoidales de aminoácidos. En el dibujo, las hélices están representadas por cilindros. La "hélice de reconocimiento" encaja con precisión en el surco mayor que recorre toda la hélice de ADN, como la rosca de un tornillo. La otra hélice alfa se ubica transversalmente al surco mayor. El patrón específico de aminoácidos presente en las hélices de reconocimiento encaja a la perfección con el patrón específico de pares de bases del sitio operador al que se une el represor.

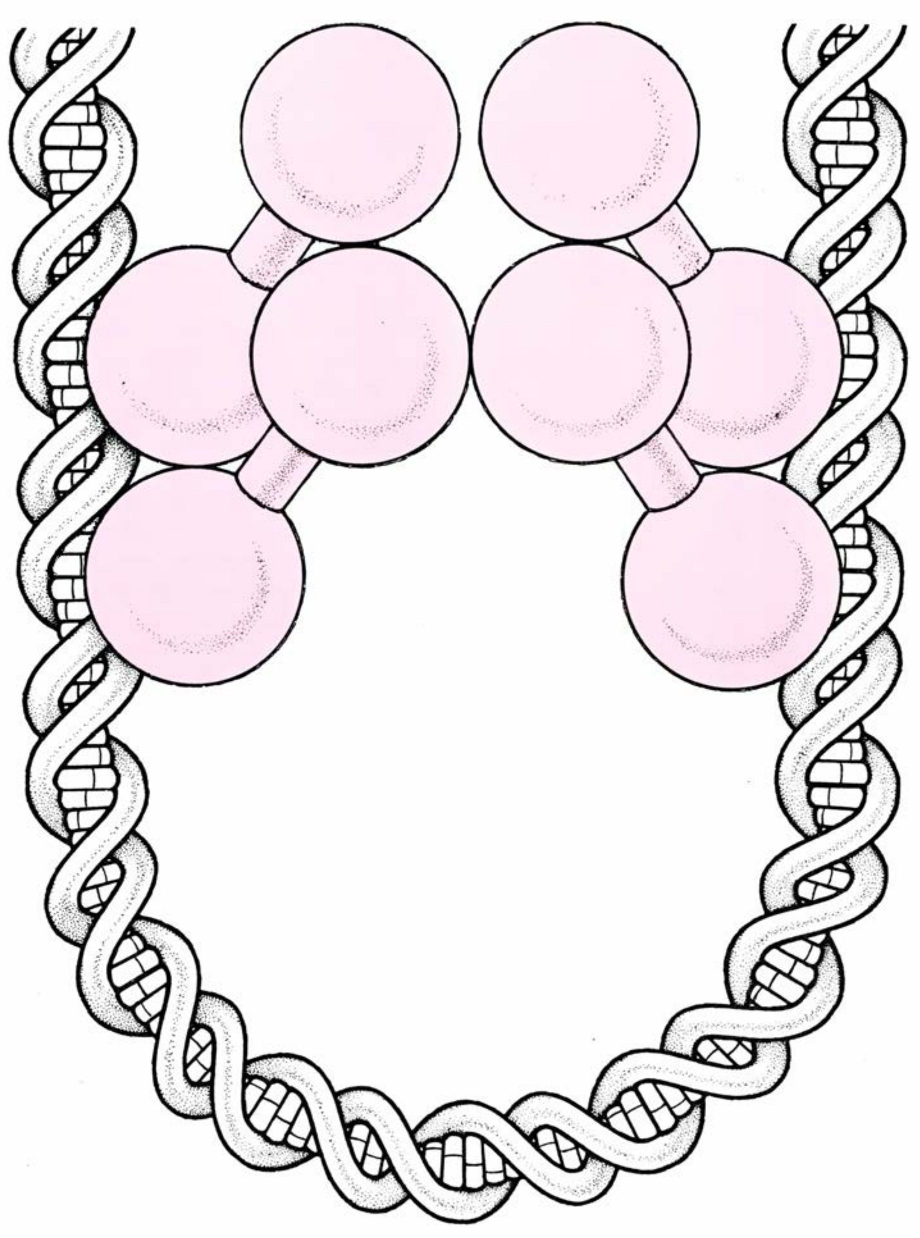
El hecho de que el represor se una en forma de dímero implica que cada una de las dos hélices alfa idénticas (en este contexto, hélices de reconocimiento) deben encontrar su sitio de encaje en la secuencia de ADN. Este requerimiento explica por qué cada sitio operador contiene dos mitades idénticas o casi idénticas, cada una de las cuales es reconocida por una hélice alfa. El dímero-represor es doblemente simétrico, lo mismo que su sitio operador.

El represor de lambda puede encontrar una secuencia específica de pares de bases entre un inmenso mar de ADN, y unirse a ella sin distorsionar gravemente la doble hélice de ADN. Muchas otras proteínas, tanto en procariotas como en eucariotas, reconocen el ADN utilizando un mecanismo similar; la secuencia de sus hélices de reconocimiento, empero, difiere de la del represor de lambda, así como las secuencias de los operadores a los que se unen. También estas hélices de reconocimiento son capaces de identificar un sitio de unión, normalmente con menos de 20 pares de bases, entre millones de secuencias distintas. (A veces, estas proteínas, y es el caso del represor, portan junto a la hélice de reconocimiento una segunda hélice alfa,

también en forma de protuberancia. En esta característica estructura "bihelicoidal", la segunda hélice alfa se sitúa transversalmente al surco mayor del ADN, coadyuvando a la orientación de la hélice de reconocimiento con respecto al surco mayor.)

Esta imagen de la interacción entre el represor y su ADN diana ha influido poderosamente en nuestra idea del posible modo de acción de los genes activadores. Cabría imaginarse, por ejemplo, que la unión del represor al ADN provocara la separación de las cadenas, y que éstas pudiesen enrollarse de nuevo adoptando una estructura insólita. Esta estructura formada por el ADN alterado podría ser la responsable de la activación del gen. Ahora bien, el hecho de que la unión del represor de lambda no cambie sustancialmente la forma del ADN sugiere que la unión *per se* al ADN no activa la transcripción.

¿Cómo activa entonces el represor la transcripción? Parte del represor, distinta de la que se une con el ADN, interacciona con la polimerasa de ARN para activar la transcripción. Esta "superficie activadora", al entrar en contacto con la polimerasa de ARN, facilita la unión de la enzima al ADN y la iniciación de la transcripción del gen del represor. No sabemos mucho sobre la estructura de la polimerasa de ARN, pero sí algo de la superficie activadora del represor.



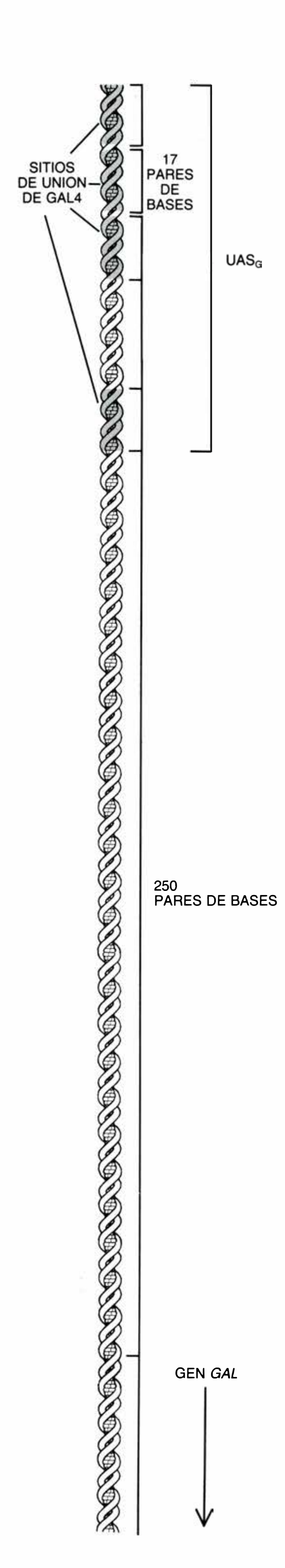
5. COOPERATIVIDAD A DISTANCIA, fenómeno que se produce cuando los sitios donde se une el represor están alejados. La formación de un bucle con el ADN que separa ambos sitios facilita el contacto entre los dímeros y la unión de los mismos con el ADN. El autor cree que la formación de lazos como éste puede facilitar también las interacciones entre proteínas implicadas en la regulación génica de organismos superiores.

Componente importante de la superficie activadora del represor de lambda es, a nuestro juicio, una hélice alfa. Basamos esta afirmación en el análisis de cierto tipo especial de represores mutantes que se unen normalmente al ADN, aunque son incapaces de activar la transcripción. Estos mutantes inducen cambios en una hélice alfa, que podemos llamar hélice de activación, o en su vecindad. El análisis de un modelo con un represor unido cerca de una polimerasa de ARN revela que la posición cabal de la hélice de activación permite un contacto perfecto con la polimerasa.

Las sustituciones de aminoácidos en los represores mutantes que se unen, aunque logran activarse, pertenecen a un tipo particular: disminuyen la cantidad de carga negativa en la hélice de activación, o cerca de la misma. La mayoría de los 20 aminoácidos que forman las proteínas son eléctricamente neutros; en la mayoría de las proteínas, los pocos aminoácidos que hay cargados positiva o negativamente aparecen en una proporción similar. Los datos de nuestros mutantes de lambda sugieren que los aminoácidos cargados negativamente revisten especial importancia para la activación, una idea a la que volveremos más tarde.

Existe otra faceta, de enorme interés, del modo de acción del represor. Tiene que ver con la unión de los dos dímeros-represores, necesarios para la activación. Los dímeros no se unen de forma independiente, sino que la unión de uno facilita la del otro. Para entender dicho efecto, imagínese el lector que un dímero-represor se une primero al punto situado más hacia la derecha (O_R1), el lugar de unión más fuerte. La presencia del represor en ese sitio ayuda a que se una otro dímero en el sitio adyacente, por contacto con él. Se dice así que los dímeros se unen al ADN cooperativamente.

Sabemos hoy que el principio de cooperatividad es moneda corriente entre las proteínas que se unen al ADN. Aunque una exposición minuciosa de ese fenómeno sobrepasa los límites del artículo, digamos, a modo de síntesis, que la cooperatividad determina que la unión de las proteínas muestre extrema sensibilidad ante los cambios en su concentración, por pequeños que sean, gracias a lo cual posibilita que la activación o inactivación de los genes se desarrolle con suma eficacia. La cooperatividad permite también que una proteína distinga con precisión el sitio del ADN donde debe unirse, entre una abundante cantidad de ADN irrelevante. En el caso del represor de lambda, el efecto cooperativo es de segundo orden: la presencia de un dímero unido



6. **GEN DE LEVADURA controlado por una secuencia activadora situada a unos 250 pares de bases "curso arriba" (*UAS_G*). Porque determina una enzima que degrada el azúcar galactosa, el gen se denomina *GAL*; la proteína reguladora que activa su transcripción es GAL4. Los cuatro sitios de unión de GAL4 presentes en la UAS_G tienen, cada uno, una longitud de 17 pares de bases, como sitio de unión del represor. La proteína GAL4 puede actuar cuando su sitio de unión está a 750 pares de bases del gen *GAL*.**

al ADN decuplica la posibilidad de que se una otro. Ahora bien, hasta este efecto resulta crítico para la acción represora.

Otro aspecto de la cooperatividad, en el caso de lambda, ha influido en nuestra idea del efecto de las proteínas reguladoras sobre la transcripción, aun cuando estén unidas a sitios muy alejados de los genes que deben controlar. Me refiero al siguiente: lo mismo que un represor puede coadyuvar a que otro se una cuando los sitios operadores están contiguos, los dímeros represores pueden ayudarse entre sí incluso en casos en los que los sitios de unión están muy alejados en el ADN. Los dímeros pueden entrar en contacto porque el ADN es flexible y crea un bucle que permite la interacción [*véase la figura 5*]. La idea del lazo fue propuesta por Sankar L. Adhya y sus colaboradores, del Instituto Nacional del Cáncer, y por Robert F. Schleif y colaboradores, de la Universidad de Brandeis.

La cooperatividad a distancia ha quedado demostrada, o rigurosamente deducida, en muchos casos de proteínas reguladoras procariotas. Lo cierto es que lambda representa una excepción, en el sentido de que sus sitios operadores, en su configuración normal, se hallan contiguos en el ADN. Es razonable suponer que el mismo mecanismo de lazo, propuesto para acercar proteínas reguladoras, podría permitir que las proteínas reguladoras unidas al ADN en un sitio determinado influyeran en la transcripción de un gen distante. Por ejemplo, un activador unido al ADN puede necesitar, para llevar a cabo su función, entrar en contacto con la propia polimerasa de ARN, o con otra proteína auxiliar enlazada en el comienzo del gen; en tal caso, la formación de un bucle en el ADN facilitaría la interacción.

¿Guarda alguna relación la forma de actuar del represor de lambda con los problemas que plantea la regulación en los organismos superiores? No era ésta una cuestión fácil de responder: cuesta más manipular eucariotas que bacterias. Aunque el desarrollo de la técnica vulgarmente conocida por "ADN recombinante" ha permitido diseñar nuevos tipos de experimentos con las células de organismos superiores, las limitaciones siguen siendo abrumadoras si las comparamos con los realizables con células bacterianas. Hemos escogido, para su estudio, los genes de levaduras, uno de los organismos eucariotas más simples. Las levaduras se multiplican con la rapidez casi de las bacterias y su manipulación genética no encierra mayor dificultad.

Nos centramos en un grupo de genes responsables de la síntesis de enzimas implicadas en la degradación del azúcar galactosa. Los genes *GAL*, que así se llaman, se activan por la proteína GAL4. La proteína GAL4 se une al ADN en una secuencia situada a unos 250 pares de bases del comienzo del más próximo de los genes *GAL*. La secuencia en cuestión es conocida con el nombre de secuencia activadora "curso arriba" (upstream) de la galactosa, o UAS_G. La expresión "curso arriba" quiere resaltar el hecho de que tal secuencia media en la actividad de GAL4 a distancia. De hecho, si se coloca la UAS_G delante o a cierta distancia de otros genes de levaduras, GAL4 activa también la transcripción de dichos genes.

Un análisis más detallado de la UAS_G revela la existencia de cuatro secuencias similares, a las que se une GAL4, cada una de ellas de 17 pares de bases. Las secuencias, como los sitios operadores de lambda, son doblemente simétricas (o casi); otros experimentos sugieren que GAL4, como el represor de lambda, se une a todos esos sitios en forma de dímero. Un dímero de GAL4 unido a uno de esos sitios está capacitado para activar la transcripción, pero no lo hará con la eficacia de varios dímeros en distintos sitios a la vez.

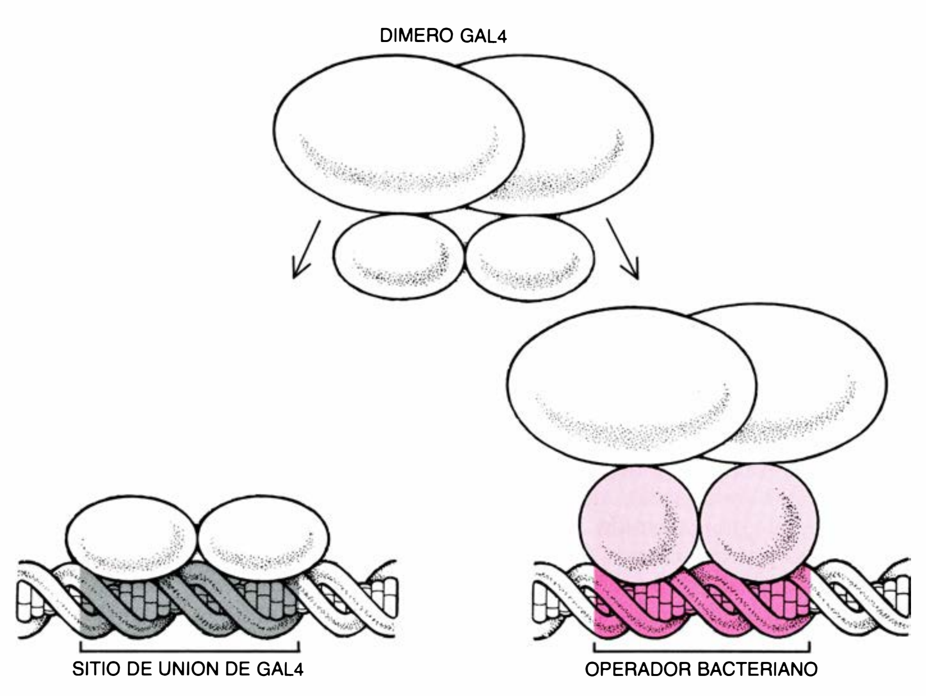
¿Cómo se comporta GAL4? Dicho de otra manera: ¿cómo se une selectivamente a la UAS_G y activa la transcripción de un gen situado a cientos de pares de bases de distancia? Los experimentos siguientes demuestran que, como en el caso del represor de lambda, GAL4 posee una superficie de unión con el ADN y una superficie activadora. Demuestran también que, a diferencia de las dos superficies del represor de lambda (muy cerca una de otra, en un mismo dominio), las regiones de unión con el ADN y activadora se encuentran en regiones distintas de la proteína y se pueden separar fácilmente; GAL4 es mucho mayor que el represor de lambda.

El experimento que permitió localizar la región de GAL4 que se une al ADN se realizó, en parte, fragmentando el gen que determina la proteína, y

volviendo a introducir los fragmentos en la levadura; se dilucidó, a continuación, qué funciones cumplían los fragmentos de proteínas sintetizados a partir de los fragmentos de los genes. Encontramos un fragmento, la parte inicial de la proteína, que se unía con el ADN, aunque no activaba la expresión génica. De este experimento sacamos la conclusión siguiente: la región activadora de GAL4 debía encontrarse en una parte de la proteína distinta de la formada por los cien primeros aminoácidos.

El experimento que confirmó la hipótesis consistió en la fusión de fragmentos de ADN del gen GAL4 con fragmentos de ADN que determinaban partes de otras proteínas fabricando así proteínas híbridas. En el experimento crítico, sustituimos la región de unión al ADN de GAL4 por la región de unión al ADN de un represor bacteriano. Según era predecible, la proteína híbrida no ejercía efecto alguno sobre la expresión génica en las levaduras normales, dada su incapacidad para reconocer la UAS_G. No obstante, si se situaba, delante de un gen de levadura, una secuencia operadora susceptible de ser reconocida por el represor bacteriano, la proteína híbrida activaba la expresión de dicho gen. El represor bacteriano nativo es incapaz de activar la expresión génica en levaduras. En otras palabras, para fabricar un activador se necesitan dos funciones: una superficie activadora (en este caso suministrada por el fragmento de GAL4) y una superficie de unión al ADN (suministrada por el represor bacteriano) para situar al fragmento activador en el ADN próximo al gen.

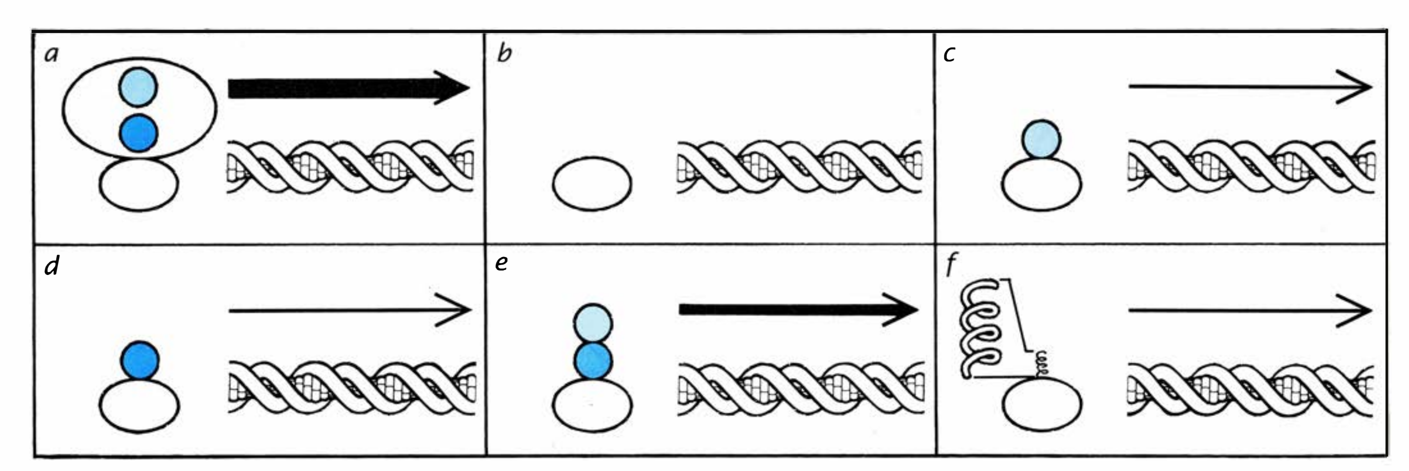
No sabemos con exactitud cómo GAL4 reconoce su sitio de unión en el ADN. A diferencia de lambda, carece de la estructura bihelicoidal característica. Su secuencia sugiere que puede constituir unas prolongaciones de aminoácidos, a manera de "dedos", sujetas en la base por átomos de zinc; de ahí que a esas estructuras se les denomine dedos de zinc. Los experimentos que se acaban de describir demuestran, no obstante, que este punto, aun siendo de gran interés, no es crucial para entender el fenómeno de la activación génica.



7. MANIPULACIONES DE GAL4 muestran que su sitio de unión con el ADN y las superficies activadoras caen en partes diferentes de la proteína. La GAL4 nativa posee dos dominios (*arriba*); una proteína GAL4 mutante que carece del dominio grande puede unirse al ADN, pero no activa la transcripción (*izquierda*). La substitución de la región de unión al ADN de GAL4 por la región de unión al ADN de una proteína bacteriana parecida al represor de lambda (*derecha*) revela que en el dominio mayor de GAL4 está la superficie activadora. Al insertar un sitio de unión para el represor bacteriano en el ADN de levadura, el híbrido activa un gen.

El experimento con los híbridos de bacteria y levadura estimuló nuestro interés por conocer más detalles sobre la superficie activadora de GAL4. El planteamiento obvio del problema venía instado por el experimento realizado con el represor de lambda: aislar mutantes de GAL4 que se uniesen al ADN, aunque privados de capacidad activadora. En el supuesto de que tales mutantes fuesen análogos a los mutantes de represor, incluirían cambios de aminoácidos que definiesen el carácter de la superficie activadora. No obstante, y por razones que se relatarán posteriormente, los mutantes que obtuvimos sólo contenían fragmentos de proteína que carecían de la mayor parte de ésta. Los mutantes, por tanto, no nos sirvieron para aproximarnos a la superficie activadora.

Intentamos una nueva aproxima-



8. DERIVADOS DE GAL4 para identificar las partes de la proteína implicadas en la activación. Cuando GAL4 está entera (*a*), dos regiones activadoras (*color*) inducen la transcripción (*flecha*); sin esas regiones no hay transcripción (*b*). Estas regiones pueden activar la transcripción independientemente, si van unidas a la región de unión al ADN de GAL4 (*c*, *d*). Juntas son casi tan eficaces como la proteína intacta, aun cuando falte casi el 80 por ciento de la proteína (*e*). Una cadena de 15 aminoácidos, que forma una hélice alfa, unida a la región de unión al ADN (*f*), funciona casi con la eficacia de las superficies activadoras.

ción, inspirada en los estudios de Keith R. Yamamoto y sus colegas, de la Universidad de California en San Francisco. Trabajaban éstos con el gen de una proteína reguladora humana, el receptor de glucocorticoides. Decidimos trocear el gen de la proteína GAL4 y recombinar la parte del gen que determina la superficie de unión al ADN con algunos de los fragmentos restantes. Pretendíamos con ello aislar una forma simplificada de GAL4, en la que la superficie activadora de dicha proteína estuviese engarzada directamente en el fragmento de unión al ADN. Encontramos dos fragmentos de la proteína, de unos 100 aminoácidos de longitud cada uno, que, por separado, activaban la expresión génica cuando llevaban a su lado el fragmento de unión al ADN. Los dos fragmentos juntos activaban la expresión génica casi tan bien como GAL4, aun cuando se había eliminado el 80 por ciento de la proteína.

Estos resultados ofrecían un doble motivo de sorpresa. Primero, el experimento funcionaba. Los biólogos moleculares acostumbran a considerar las proteínas estructuras tridimensionales muy bien definidas, que no se alteran fácilmente. Una cosa es (al menos así pensamos nosotros) pegar trozos de proteínas entre sí, y esperar que el híbrido funcione, y, otra distinta, eliminar grandes segmentos de una proteína, y esperar que siga manteniendo su actividad.

Quizá no hubiera razón para dicha sorpresa. En los últimos años se han ido acumulando pruebas de que los dominios de las proteínas pueden ser unidades funcionales independientes. Aunque la de GAL4 sea una estructura todavía por dilucidar, el hecho de que las regiones activadoras y la función de unión al ADN puedan separarse y recombinarse indica que cada región de éstas se halla en un dominio distinto.

El segundo motivo de sorpresa de nuestro análisis de las regiones activadoras de GAL4 sobreviene al comparar sus secuencias. Es frecuente que segmentos de proteínas que realizan funciones similares generen secuencias y estructuras tridimensionales similares. En este caso, sin embargo, sólo observamos un elemento común: la presencia de un gran exceso de aminoácidos cargados negativamente. Kevin Struhl y sus colegas, de la facultad de medicina de Harvard, habían demostrado que la región activadora de otro activador transcripcional de levaduras, el GCN4, contenía un exceso de cargas negativas; por lo demás, su secuencia no guardaba mucha relación con ninguna de las regiones activadoras de GAL4.

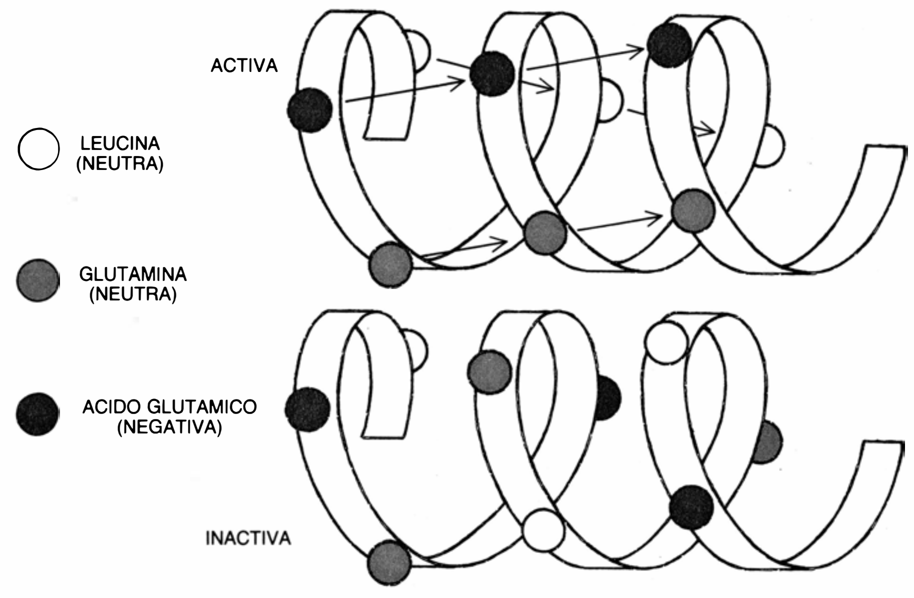
Con extraordinaria facilidad se hallan secuencias de proteínas que actúan de secuencias activadoras si van unidas a un fragmento de proteína trabada con el ADN. Nosotros recombinamos trozos aleatorios de ADN bacteriano con la parte del gen de la proteína GAL4 que determinaba la región de unión al ADN. Introdujimos, a continuación, estos genes híbridos en levaduras; queríamos averiguar si las proteínas híbridas podían activar a un gen control, junto al cual había un sitio de unión de GAL4.

Un alto porcentaje de las nuevas proteínas híbridas activaban la expresión del gen. Sus secuencias no se parecían, salvo en que portaban todas ellas un exceso de aminoácidos cargados negativamente. Algunos de los nuevos activadores, cuyas regiones activadoras poseían sólo unos 50 aminoácidos, alcanzaban la eficacia de la proteína GAL4 entera.

El exceso de carga negativa aludido en tantos activadores da a entender que no se trata de un fenómeno casual. Lo corrobora, además, un nuevo experimento genético. Lo dijimos antes, en los intentos realizados para definir la región activadora de GAL4, mediante el aislamiento de mutantes específicamente afectados en la función activadora, sólo se obtuvieron fragmentos de la proteína. Si en GAL4 hay dos regiones activadoras, y pueden obrar por separado, para perder totalmente la función activadora un mutante tendría que sufrir cambios que afectasen al menos a dos partes de la proteína, acontecimiento que podría ser de prohibitiva rareza. Tan sólo en el supuesto de que se perdiese la mayor parte de la proteína se eliminarían las dos regiones activadoras.

Comenzando el experimento, sin embargo, con una versión simplificada de GAL4, se pueden aislar proteínas mutantes que activen con eficacia pareja a la mostrada por la proteína nativa. Se observa una buena correlación (aunque no perfecta) entre carga y actividad: los mutantes con mayor actividad suelen llevar sustituciones de aminoácidos que incrementan la carga negativa, mientras que los mutantes con menor actividad portan generalmente sustituciones que disminuyen dicha carga.

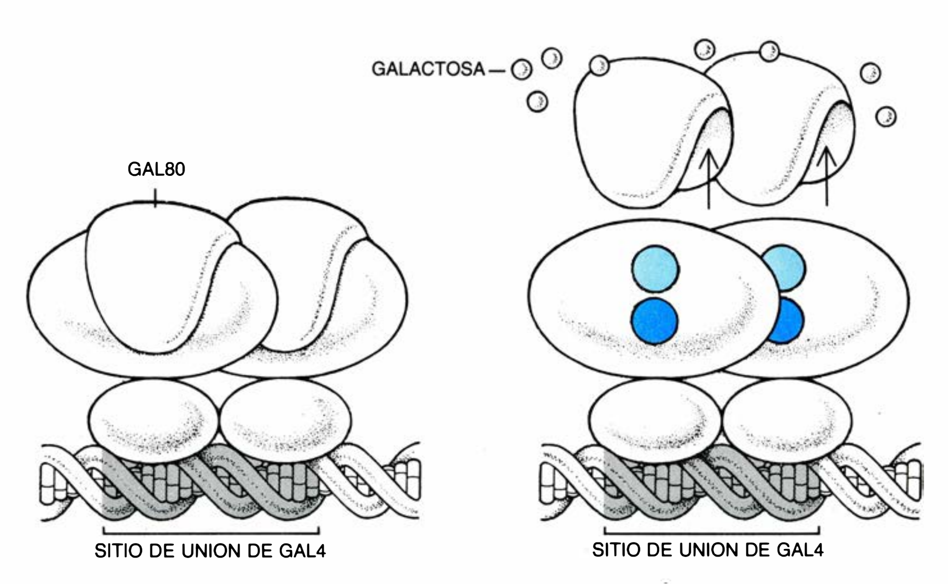
Algunas excepciones sugieren, sin embargo, que además de la cantidad de carga negativa que lleva la superficie activadora, debe intervenir algún aspecto estructural. Inspirados por el ejemplo de la superficie activadora del represor de lambda, decidimos probar si una hélice alfa con cargas negativas en una de sus superficies tenía capacidad activadora. Diseñamos un fragmento de ADN con información para la síntesis de un pequeño segmento de proteína que, en teoría, podía plegarse adoptando una configuración en hélice alfa. A dicho fragmento de ADN le adosamos el ADN que determinaba la parte de unión al ADN de GAL4. Si se llegaba a formar la hélice alfa, llevaría no sólo una superficie cargada negati-



9. CADENAS DE AMINOACIDOS SINTETICAS fijadas a la región de unión con el ADN de la proteína GAL4 para estudiar los requerimientos necesarios en una superficie activadora. En ambas cadenas aparecen los mismos aminoácidos, aunque en diferente orden. La disposición de los aminoácidos se hizo de tal manera que, si las cadenas formasen una hélice alfa (como se muestra en el dibujo), una de ellas tuviese alineados los aminoácidos cargados negativamente en una de sus superficies (*arriba*) y la otra los mostrara dispersos (*abajo*). La cadena con las cargas alineadas activa la transcripción; no así la otra. Por consiguiente, la función activadora dependerá de la distribución de la carga, más que de la carga propiamente dicha.

vamente, sino además una superficie hidrofóbica, que quizá podría ayudarle a empaquetarse sobre la superficie del fragmento GAL4.

Esta proteína híbrida, con una superficie activadora constituida por sólo 15 aminoácidos, activaba genes en levaduras con una eficacia aproximada del 20 por ciento con respecto a la proteína GAL4 intacta. Ahora bien, si se trastocaba el orden de los aminoácidos, el fragmento de proteína perdía su capacidad activadora, mostrando así que la carga por sí sola no confería capacidad activadora. Probablemente, para crear una región activadora se requiere cierta distribución de cargas negativas en una superficie de la hélice.



10. LA GALACTOSA CONTROLA A GAL4 a través de una proteína denominada GAL80. Cuando no hay galactosa en el medio donde crece la levadura, GAL80 encubre las superficies activadoras de GAL4 e impide así que la proteína active la transcripción. En presencia de galactosa, el azúcar, o uno de sus derivados metabólicos, separa GAL80, quedando GAL4 libre para activar a los genes de las enzimas que degradan el azúcar.

Los experimentos que se han descrito sentaron las bases de una nueva generalización que muchos de nosotros no esperábamos: GAL4, y los fragmentos de GAL4 que activan la transcripción en levaduras, lo hacen también en células de mamíferos, en células de *Drosophila*, la mosca de la fruta, e incluso en células de plantas de tabaco. Los experimentos realizados para demostrarlo siguieron una estrategia común. Para ello se necesitaba introducir el gen de la proteína GAL4, o uno de sus derivados, en una célula foránea, que es la que sintetiza la proteína. A continuación, se insertaba un sitio de unión de GAL4 cerca de uno de los genes nativos del propio organismo. En todos los casos —plantas, insectos y mamíferos—, el activador de levaduras incrementaba la transcripción del gen nativo.

No observábamos activación cuando se omitía el sitio de unión de GAL4 o cuando se expresaba la región de unión al ADN de GAL4 sin una región activadora. En definitiva, una molécula dotada de una región activadora junto a un fragmento de unión a ADN activará la expresión génica en muchos, si no en todos, los eucariotas.

Dado que una región activadora conserva su función en células foráneas, ¿conserva también, en dichas células, su capacidad para actuar a grandes distancias? En los genes de organismos superiores se encuentran, con bastante frecuencia, sitios de unión a una distancia bastante considerable, sitios de unión que reciben el nombre de intensificadores ("enhancers"). ¿Pueden nuestros activadores estimular la transcripción aunque estén unidos al ADN a más de 1000 pares de bases del gen? La respuesta es afirmativa, siempre que el activador abarque una región activadora suficientemente potente. Una región con esas características se encuentra en una proteína reguladora, la VP16, sintetizada por el virus herpes simplex. Una proteína híbrida, fruto de la fusión entre VP16 y el fragmento de unión al ADN de GAL4, puede, en células de mamíferos, activar la transcripción de un gen situado a más de 1000 pares de bases "curso arriba" o "curso abajo" de una UAS_G.

Nos parece que este segmento de VP16 posee la combinación de estructura y carga adecuada para interaccionar, de manera especialmente intensa, con su proteína diana. GAL4 interviene sólo a distancias más cortas, presumiblemente porque su región activadora interacciona de forma menos enérgica con la proteína diana.

Hasta el momento, mi análisis de la regulación génica ha soslayado la cuestión del control de las propias proteínas reguladoras, que posibilita que una célula responda como es debido a las condiciones ambientales. En parte al menos sí sabemos cómo se produce el control de GAL4. Esta proteína pone en marcha los genes que degradan el azúcar galactosa; importa, pues, que GAL4 actúe cuando hay galactosa en el medio donde se encuentran las levaduras, y no lo haga en ausencia del azúcar. Se observa que GAL80, una proteína inhibidora, protege normalmente las superficies activadoras de GAL4; en presencia de galactosa, el azúcar, o uno de sus productos metabólicos, suelta el inhibidor, dejando expuestas las superficies activadoras.

Los estudios y descubrimientos que se han presentado platean problemas no resueltos. En primer lugar, si aceptamos que las regiones activadoras interaccionan con otra proteína implicada en la transcripción, ¿cuál es esa otra proteína? Podría tratarse de la propia polimerasa del ARN, aunque sospechamos que será otra proteína la que actúe de intermediario entre la proteína reguladora y la polimerasa. Evidentemente, una proteína de ese tipo debe estar presente en muchos eucariotas (si no en todos).

En segundo lugar, ¿de qué naturaleza es la interacción entre la región activadora y su diana? Los biólogos moleculares suelen admitir que las interacciones proteína-proteína están determinadas por las características estructurales específicas de las moléculas que intervienen. En este caso, empero, moléculas cuyas secuencias cargadas negativamente son muy distintas funcionan a pleno rendimiento. Conviene, por último, señalar que los activadores que hemos estudiado podrían constituir sólo una entre las muchas clases de proteínas activadoras existentes.

BIBLIOGRAFIA COMPLEMENTARIA

AN OPERATOR AT −280 BASE PAIRS THAT IS REQUIRED FOR REPRESSION OF *ARA-BAD* OPERON PROMOTER: ADDITION OF DNA HELICAL TURNS BETWEEN THE OPERATOR AND PROMOTER CYCLICALLY HINDERS REPRESSION. Teresa M. Dunn, Steven Hahn, Sharon Ogden y Robert F. Schleif en *Proceedings of the National Academy of Sciences of the United States of America*, vol. 81, n.° 16, págs. 5017-5020; agosto de 1984.

A GENETIC SWITCH: GENE CONTROL AND PHAGE λ. Mark Ptashne. Cell Press & Blackwell Scientific Publications, 1987.

HOW EUKARYOTIC TRANSCRIPTIONAL ACTIVATORS WORK. Mark Ptashne en *Nature*, vol. 335, n.° 6192, págs. 683-699; 20 de octubre de 1988.

Terremotos profundos

Han constituido un fructífero enigma desde su descubrimiento, hace 60 años. ¿Cómo pueden romperse las rocas a las temperaturas y presiones reinantes a centenares de kilómetros de profundidad?

Cliff Frohlich

En la mayoría de terremotos, la corteza terrestre se rompe como la porcelana. Se acumulan las tensiones hasta que se desarrolla una fractura a una profundidad de algunos kilómetros y el desplazamiento las libera. Sin embargo, algunos terremotos tienen lugar donde la tierra no puede fracturarse. Se producen a centenares de kilómetros, en el interior del manto, donde la alta presión, se cree, impide que la roca se rompa, a pesar de hallarse sometida a tensiones capaces de deformarla como si fuera masilla. ¿Por qué hay terremotos a tales profundidades?

Estos episodios profundos y misteriosos abundan. Desde 1964, el Centro Sismológico Internacional de Londres (CSI) ha catalogado más de 60.000 terremotos de profundidades superiores a los 70 kilómetros: el 22 por ciento de todos los terremotos localizados en ese intervalo. Pocos han producido daños. En su inmensa mayoría, los terremotos catastróficos son someros, ya que ocurren a 50 kilómetros de profundidad o menos; aunque no se olvide que un temblor centrado a una profundidad de 100 kilómetros devastó Bucarest, la capital de Rumanía, el 4 de marzo de 1977.

Otra cosa es su repercusión geofísica, que, en la mayoría de esos acontecimientos profundos, resulta ser máxima. Su patrón geográfico aportó indicios para la gran teoría unificadora de la geofísica moderna, la tectónica de placas. También se han revelado como fuentes energéticas adecuadas para el estudio sísmico del interior profundo de la Tierra, que trata de descifrar su estructura a partir del comportamiento de las ondas sísmicas en su propagación. Los terremotos profundos pueden estar descubriéndonos sus propios secretos. Las observaciones sismológicas, junto con estudios de laboratorio del comportamiento de las rocas a altas presiones, han conducido a explicaciones verosímiles sobre el modo en que pueden ocurrir.

CLIFF FROHLICH se halla adscrito al Instituto de Geofísica de la Universidad de Texas en Austin desde 1978, dos años después de que la de Cornell le concediera el doctorado. Durante la elaboración de la tesis estudió la estructura del manto que yace bajo el Pacífico sudoccidental mediante el análisis de las ondas sísmicas que proceden de terremotos profundos. Desde entonces, ese capítulo de la geofísica constituye el centro de su investigación.

La profundidad de todos los terremotos fue motivo de controversia durante las décadas que precedieron al año 1927, cuando el sismólogo japonés Kiyoo Wadati demostró convincentemente la existencia de episodios profundos. Percatándose de que, en algunos casos, la sacudida intensa quedaba confinada en un dominio restringido, algunos autores habían supuesto que las fuentes del terremoto debían hallarse a pocos kilómetros de la superficie. Otros llegaron a defender profundidades focales mucho mayores, de hasta 1200 kilómetros, al tratar de determinar direcciones sísmicas a partir de las desviaciones que sus ondas producían en los sismógrafos elementales de la época.

La controversia se encrespó con el mejor conocimiento de las ondas sísmicas y su propagación. Mediante el estudio de sismogramas, se aprendió a identificar varios tipos de ondas internas (ondas sísmicas que se propagan por el interior de la Tierra, a diferencia de las ondas superficiales, que lo hacen por su superficie). Las primeras en llegar son las ondas *P*, u ondas primarias. Llamadas también ondas de compresión, son ondas sonoras en la Tierra, que se propagan como zonas alternantes de presión alta y baja. Les siguen las ondas *S* —ondas secundarias o de cizalla— que se propagan a través de la Tierra convertidas en sacudidas de un lado a otro. Comparaciones de los instantes de llegada a diferentes estaciones de las ondas *P* y *S* de un terremoto determinado mostraron que el tiempo de propagación dependía de la distancia y de la estructura interna del planeta.

En teoría, dado un modelo de la estructura terrestre y un conjunto de tiempos de llegada registrados en varias localidades, debería poderse establecer el lugar del acontecimiento. En 1922, H. H. Turner, director del Centro de Recogida de Datos Sismológicos que luego se convirtió en el CSI, aplicó este método en un estimulante y polémico artículo. Basándose en un análisis de datos de estaciones de todo el mundo, Turner sugirió que los terremotos ocurrían en tres gamas de profundidades. Los episodios de "foco alto" tenían su fuente cerca de la superficie, pero los terremotos normales, los más abundantes, se producían a unos 150 kilómetros. Los de "foco profundo" sucedían a profundidades de hasta 650 kilómetros.

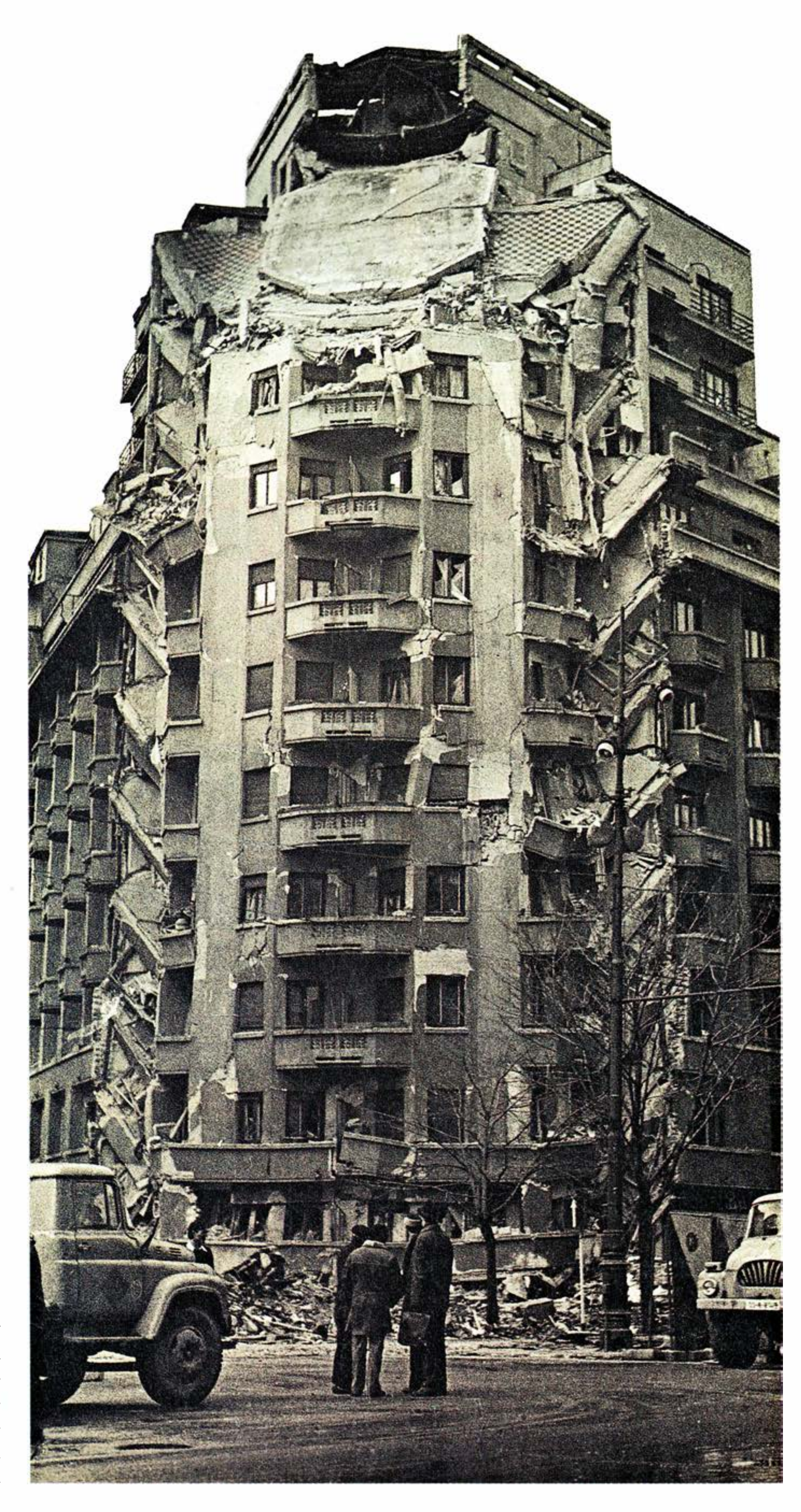
El método de Turner era correcto, pero sus datos estaban poco elaborados para los planteamientos modernos, su conocimiento de la estructura del interior profundo de la Tierra era incompleto y los imprecisos relojes de la época provocaron errores de segundos, e incluso minutos, en los tiempos de llegada de las ondas. Turner convenció a muy pocos contemporáneos. S. K. Banerji, del Observatorio de Bombay, señaló que si la mayoría de terremotos tenían su foco a 150 kilómetros, o más, de profundidad, serían contados los episodios que produjeran ondas superficiales fuertes, y, sin embargo, el propio catálogo de Turner daba cuenta de abundantes ondas superficiales. Harold Jeffreys, de la Universidad de Cambridge, planteó una objeción más seria: negaba, lisa y llanamente, que hubiera

terremotos a las profundidades propuestas por Turner.

Por debajo de los 50 kilómetros, sostenía Jeffreys, el calor y la presión cambian la roca del manto; lo que era un material quebradizo y capaz de fracturarse lo truecan en otro dúctil. En apoyo de su argumentación, Jeffreys señaló que desde el fin de la glaciación más reciente, las costas de Canadá y Europa septentrional, liberadas de la carga de los glaciares, habían ascendido como si el manto subyacente fuera capaz de fluir bajo esfuerzo, como un líquido muy viscoso. Además, adujo experimentos de laboratorio que confirmaban que, a temperaturas y presiones elevadas, las rocas se deformaban gradualmente en respuesta a un esfuerzo aplicado, en vez de fracturarse súbitamente.

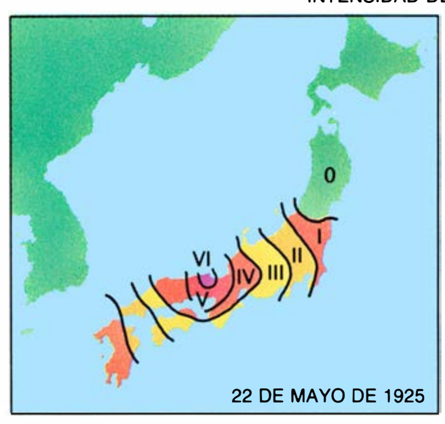
Kiyoo Wadati, joven empleado de la Agencia Meteorológica Japonesa, no refutó el argumento de Jeffreys; se limitó a presentar pruebas convincentes de la profundidad notable de algunos temblores. La frecuencia y la potencia destructora de los terremotos en Japón habían llevado al gobierno japonés a establecer lo que entonces era la mejor red de estaciones sismográficas del mundo. Por consiguiente, Wadati disponía de abundantes datos y aplicó nuevos métodos para determinar las profundidades de los terremotos. En lugar de comparar los instantes de llegada a diferentes estaciones sismográficas, como hizo Turner, se basó en una diferencia de tiempo que podía medirse fácilmente en cada estación, aunque los relojes fueran imprecisos: el intervalo entre la llegada de las ondas *P* y la de las ondas *S*. Dado que cada clase de ondas se propaga a una velocidad constante, el intervalo aumenta en proporción a la distancia entre la estación y el foco del terremoto.

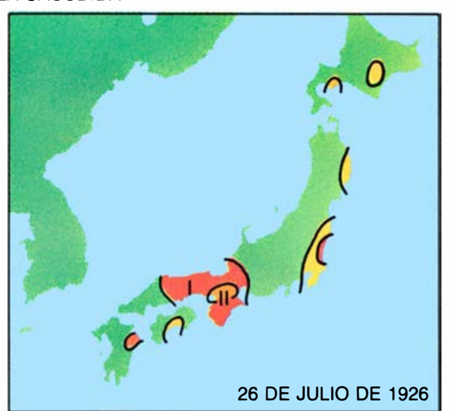
Wadati descubrió que, para la mayoría de terremotos, el intervalo era muy pequeño cerca del epicentro, el punto donde la sacudida es máxima. Sin embargo, para algunos episodios, el intervalo era largo incluso en el epicentro. Wadati halló un patrón similar cuando analizó datos de la intensidad de la sacudida. La mayoría de terremotos tenía un área pequeña de vibración intensa que se debilitaba rápidamente al incrementarse la distancia al epicentro, pero otros se caracterizaban por un pico menor de la intensidad, sentido en una extensión mayor. Los intervalos *P-S* y los patrones de intensidad sugirieron dos clases de terremotos: los someros, en los que el foco

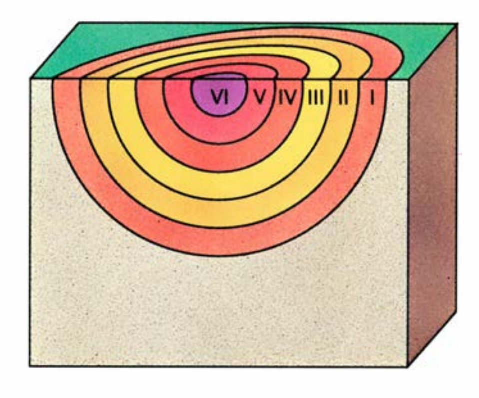


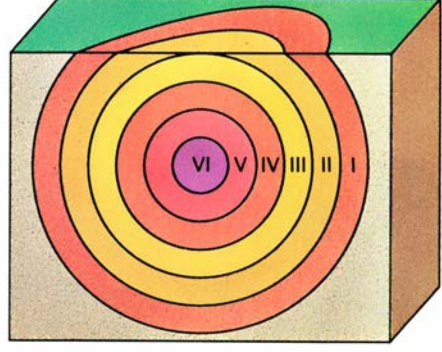
1. EL TERREMOTO PROFUNDO DESTRUCTIVO se inscribe entre los acontecimientos raros. El único grave de las últimas décadas tuvo lugar el 4 de marzo de 1977, a 150 kilómetros de la superficie de Bucarest. Las columnas de la planta baja de este edificio se derrumbaron, lo que determinó que el chaflán se rebajara en un piso. Los daños personales que el episodio produjo en la capital rumana se cobraron 1500 víctimas. La fotografía se debe a Neculai Mandrescu, adscrito al Instituto de Física de la Tierra de Bucarest.

INTENSIDAD DE LA SACUDIDA

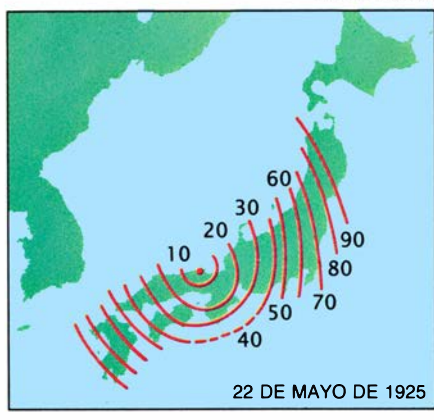


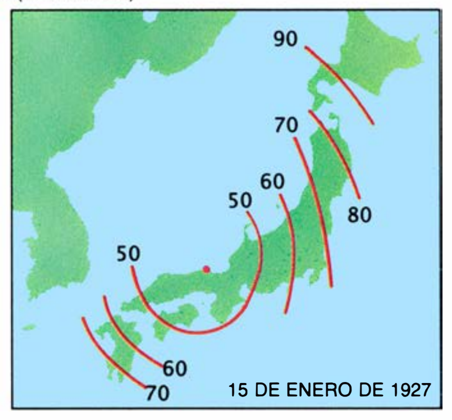


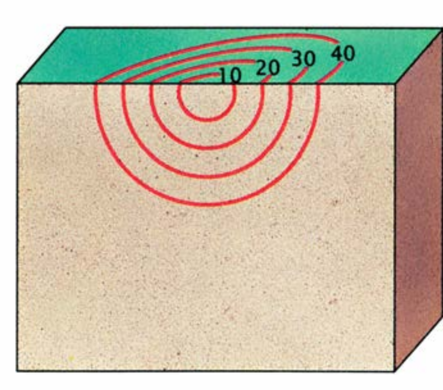


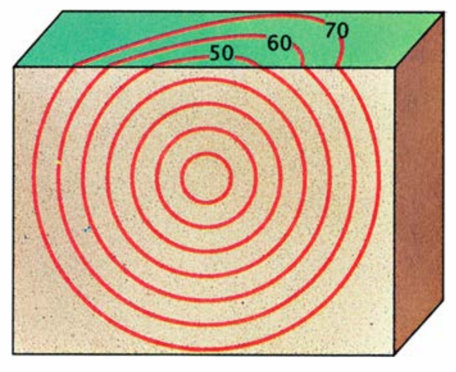


DIFERENCIA S-P (SEGUNDOS)









2. A TRAVES DE LAS MANIFESTACIONES SUPERFICIALES distinguimos los episodios someros de los profundos. El mapa compara los datos de la intensidad de la sacudida y del intervalo entre la llegada de las ondas *P* (de compresión) y las *S* (de cizalla) de varios terremotos registrados entre los años 1925 y 1927. Reunió los datos el sismólogo japonés Kiyoo Wadati. La sacudida fue muy intensa cerca del epicentro de un episodio de 1925, pero decayó rápidamente con la distancia (*arriba, a la izquierda*); un terremoto de 1926 produjo una sacudida menos intensa que se desvaneció más lentamente (*arriba, a la derecha*). La diferencia *S-P* en el epicentro del primer episodio no llegó a 10 segundos, pero el intervalo aumentó rápidamente con la distancia (*abajo, a la izquierda*); la diferencia *S-P* del episodio de 1927 tuvo un mínimo más largo —unos 40 segundos—, pero aumentó más lentamente (*abajo, a la derecha*). De los datos, Wadati sacó la conclusión de que el terremoto de 1925 tuvo su foco cercano a la superficie y los episodios posteriores se registraron a 400 kilómetros de profundidad, aproximadamente (*cortes*). (Los dibujos han sido realizados por Joe Lertola.)

subyace inmediatamente bajo el epicentro, y los profundos, con un foco a varios cientos de kilómetros.

Otros investigadores aplicaron las técnicas de Wadati a datos de terremotos de otras áreas geográficas y confirmaron los resultados de éste: los terremotos "normales" tenían una profundidad focal de 50 kilómetros o menos, mientras que algunos episodios arrancaban de mucho más abajo: de hasta 600 kilómetros o más. Turner se había equivocado respecto a la profundidad de los terremotos normales, pero los episodios profundos existían. También Banerji había estado en lo cierto: los sismogramas de episodios cuya profundidad se había confirmado mostraban que las ondas superficiales producidas eran débiles.

¿Qué decir sobre la afirmación hecha por Jeffreys de que la roca del manto situada a más de 50 kilómetros de profundidad era demasiado dúctil para almacenar el esfuerzo necesario para un terremoto? Una observación de Wadati anunciaba parte de la respuesta:los terremotos profundos no ocurren en roca de manto ordinaria. En 1935, Wadati publicó un mapa en el cual mostraba la ubicación de los terremotos desencadenados cerca de Japón y sus profundidades focales. Había encontrado que los focos se localizaban a lo largo de curvas de nivel aproximadamente paralelas, y que las profundidades de las curvas aumentaban continuamente desde la costa oriental de Japón hacia el oeste. Comentó: "La posibilidad de trazar curvas de nivel de la profundidad focal sugiere que, en la corteza, hay algo así como una *superficie débil*..., donde probablemente se desencadena el terremoto. Esta superficie se extiende como un talud en la corteza cerca de las islas japonesas".

Las profundidades sísmicas de otras regiones, proseguía Wadati, definían superficies inclinadas parecidas: "Los terremotos de foco profundo son aptos para producirse en el lado más próximo al continente, y los de foco somero, en el otro [el cual], en la mayoría de casos, está rodeado por un mar muy profundo. Esta tendencia parece observarse en muchas regiones volcánicas del mundo".

Casi todos los terremotos profundos se ajustan al patrón que describió Wadati. Dondequiera que abundan —generalmente en el borde de un océano profundo— definen una zona inclinada que se extiende desde cerca de la superficie hasta una profundidad de 600 kilómetros o más. Hoy conocidos como zonas de Wadati-Benioff, en honor de

Wadati y del sismólogo Hugo Benioff, que cartografió estas zonas en los años cuarenta y cincuenta, los patrones de terremotos proporcionaron pruebas cruciales para el nuevo paradigma geofísico que emergió en los años sesenta.

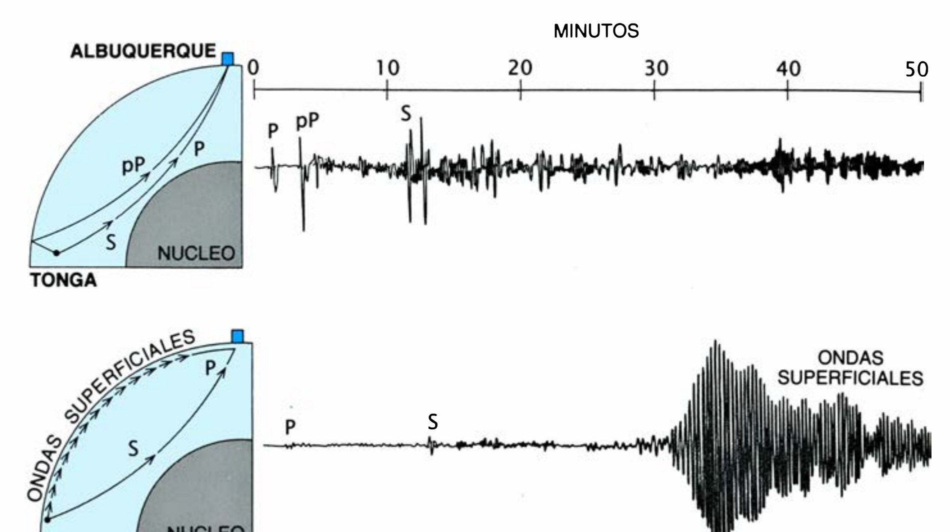
Esta nueva concepción acerca de la Tierra explica muchos de los grandes accidentes superficiales del planeta y buena parte de la historia geológica en términos de un conjunto de placas móviles que cubren la superficie terrestre. Las placas se separan en las dorsales mesoceánicas; allí donde chocan, generalmente en el límite de cuencas oceánicas, levantan montañas y esculpen los márgenes de los continentes. Este proceso de tectónica de placas constituye la expresión superficial de la convección, o circulación dirigida por calor, en el manto terrestre. El material caliente asciende desde el interior del manto y circula horizontalmente cerca de la superficie terrestre. El techo del flujo horizontal, hasta un grosor de unos 50 kilómetros, se enfría y forma las placas rígidas, que incluyen la corteza terrestre y parte del manto subyacente.

El flanco frío descendente del circuito se halla donde las placas convergen. Allí, una placa se subduce, es decir, se dobla bajo la otra placa y vuelve a caer al manto. Los terremotos profundos ayudaron a establecer la realidad de la subducción cuando se comprobó que tenían lugar en una laja descendente y que la zona de Wadati-Benioff dibujaba su forma. La profunda fosa que suele haber junto a la zona de Wadati-Benioff, en su lado oceánico —el "mar profundo" de la descripción de Wadati—, señala el encorvamiento hacia abajo de la placa subducida; la línea de volcanes que a menudo se forma en las inmediaciones se alimenta de material fundido que asciende desde la laja. Wadati acertó al anticiparse: en su artículo de 1935, había planteado la hipótesis de que los terremotos y volcanes de las inmediaciones de Japón debían ser el resultado de deriva continental (la teoría precursora de la tectónica de placas), propuesta veinte años antes por Alfred Wegener.

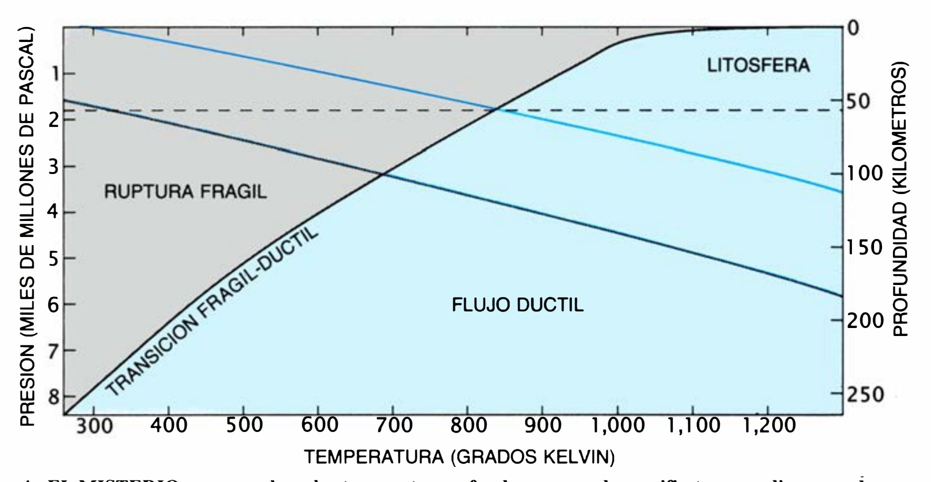
Ese esquema geofísico proporciona parte de la respuesta a la objeción de Jeffreys. Los terremotos profundos que se desencadenan en una zona de Wadati-Benioff se producen en unas rocas que están centenares de grados más frías que el manto circundante y, por tanto, son menos dúctiles y se hallan más capacitadas para almacenar energía elástica. Pero otros factores parecen afectar también a la distribución de terremotos profundos. Sus focos, por ejemplo, no están uniformemente repartidos a lo largo de la zona de Wadati-Benioff, sino que los cambios en la frecuencia de los terremotos parecen coincidir con profundidades donde la estructura cristalina de las rocas del manto cambia a una fase más densa, como resultado de la presión creciente.

A unos 400 kilómetros, por ejemplo, los estudios sismográficos muestran un incremento súbito de la velocidad de las ondas sísmicas, lo que indica un aumento de densidad de la roca. Allí, el olivino —un compuesto constituido por silicatos con proporciones variables de hierro y magnesio, que es el principal integrante del manto y de la placa que se subduce— cambia a una fase cristalina más densa, conocida como espinela. En las inmediaciones de esa región, el número de terremotos profundos se reduce al mínimo.

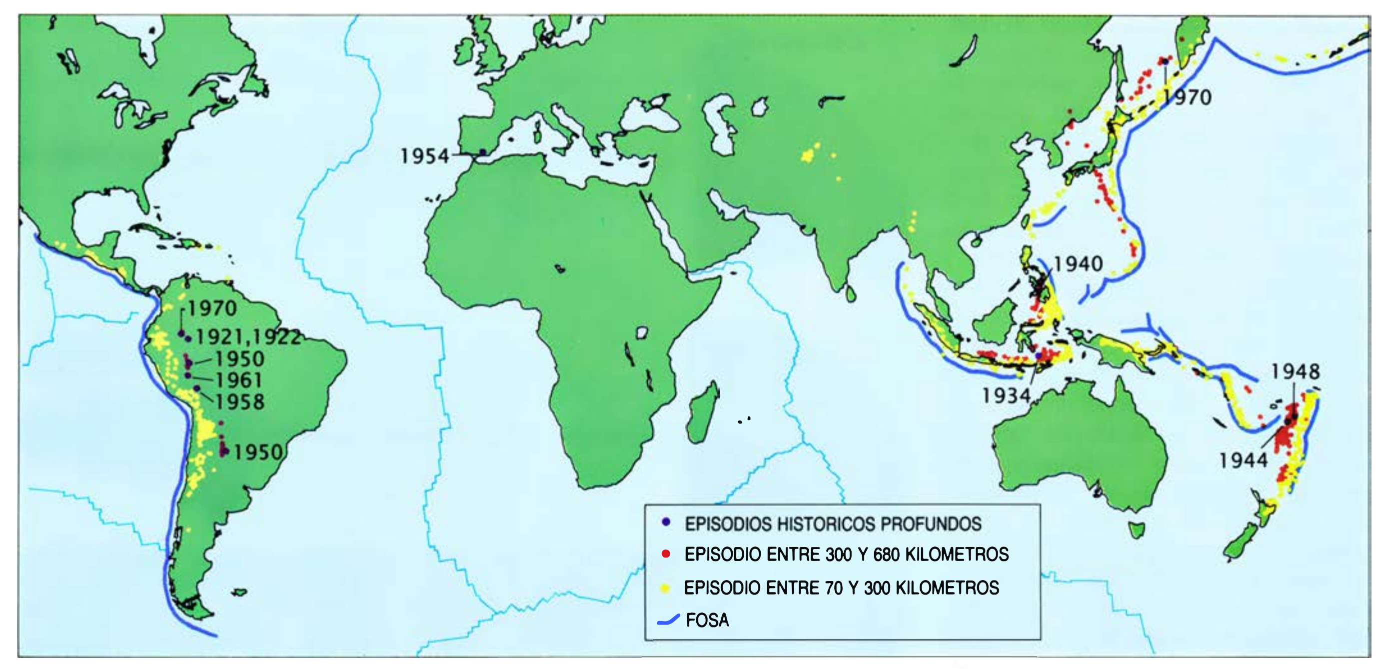
Algunas zonas de subducción son tranquilas por debajo de esta transición primera. Pero las zonas que a estas profundidades experimentan terremotos muestran su máximo nivel de actividad localizado más abajo, entre la transición citada y otro límite, menos conocido, situado a unos 650 kilómetros de profundidad. De nuevo, un acusado aumento de velocidad de las ondas sísmicas señala el límite; discrepan, sin embargo, los investigadores sobre la interpretación del incremento de densi-



3. SISMOGRAMAS DE TERREMOTOS PROFUNDOS Y SOMEROS de Tonga se registraron en Albuquerque, Nueva México, a casi un cuadrante de la vuelta al mundo. El acontecimiento profundo (*arriba*) produjo ondas *P* y *S* fuertes, que atravesaron la tierra a diferentes profundidades. Algunas ondas *P* adquirieron la forma de ondas *pP*, que se propagan hacia arriba desde el foco hasta la superficie terrestre y se reflejan de vuelta al interior. A causa de su profundidad —unos 625 kilómetros— el episodio produjo pocas ondas superficiales. Las ondas *P* y *S* del episodio somero (*abajo*) fueron relativamente débiles; la mayor parte de la energía se observó a través de ondas superficiales, la última forma de las señales que llegaron.



4. EL MISTERIO que envuelve a los terremotos profundos se pone de manifiesto en un diagrama de condiciones de presión y temperatura bajo las cuales la roca cambia desde una condición quebradiza, que puede romperse bajo esfuerzos y producir un terremoto, a un medio dúctil, que responde a los esfuerzos deformándose gradualmente. La presión y la temperatura aumentan con la profundidad, de modo que la roca suele volverse dúctil a unos 60 kilómetros bajo la superficie (*curva azul*). La roca de zonas de terremotos profundos está anómalamente fría, pero aunque estuviera 500 grados más fría que el manto "normal", sería dúctil a unos 100 kilómetros (*curva gris*). Ello no obstante, se han registrado terremotos a casi 700 km.



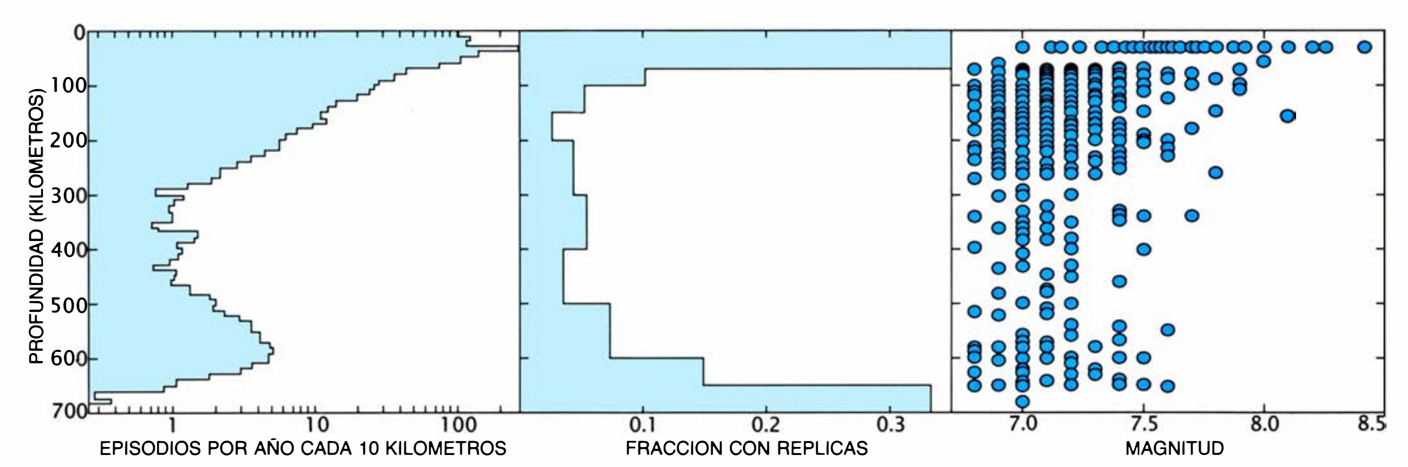
5. MAPA DE LOS EPISODIOS PROFUNDOS registrados durante los últimos 25 años, jerarquizados por la hondura; también muestra terremotos históricos muy grandes (los de magnitud superior a 7) a profundidades extremas (más de 630 kilómetros), con el año en que ocurrieron. Casi todos los terremotos profundos se producen cerca de una fosa oceánica, donde una de las placas rígidas de la litosfera —la corteza y el manto— sufre un proceso de subducción.

dad que aquél indica: ¿representa un segundo cambio de fase o un cambio de composición? En cualquier caso, la actividad sísmica cae súbitamente a cero cerca de este segundo límite.

En un esfuerzo por determinar la máxima profundidad de los eventos profundos, Philip B. Stark, entonces en la Universidad de Texas en Austin, y yo aplicamos diversos procedimientos de análisis. Examinamos, entre otras cosas, los lapsos que se dan entre las ondas ordinarias *P* y las *pP* (ondas de presión que se propagan hacia arriba, hacia la superficie, se reflejan y regresan por el interior terrestre hasta una estación sismográfica distante). Descubrimos que los episodios más profundos se originaron a 680 o 690 kilómetros. Beth A. Rees y Emile A. Okal, de la Universidad del Noroeste, realizaron un análisis similar, con resultados parecidos.

La desaparición de la actividad sísmica, demasiado brusca, no puede deberse al reblandecimiento gradual de la laja por calentamiento desde el manto circundante. Quizás indique que las lajas subducidas no pueden trascender el límite de los 650 kilómetros. En tal caso, la convección puede quedar confinada al manto superior, situado por encima del límite, mientras que el material del manto superior tal vez no se mezcle nunca con el del inferior. En cuyo caso, el cambio de velocidad de las ondas sísmicas a 650 kilómetros marca, probablemente, un cambio de composición del manto. Pero cabe una posibilidad alternativa (planteamiento que ha provocado una de las más acaloradas controversias en geofísica de la parte sólida de la Tierra): la laja subducida podría penetrar en el manto inferior. En cuya hipótesis, la convección involucraría todo el manto y el límite de los 650 kilómetros señalaría un cambio de fase en un medio de composición uniforme. Tal cambio correspondiente en las propiedades mecánicas de la roca sería responsable de la inexistencia de actividad sísmica.

El descenso de una placa subducida



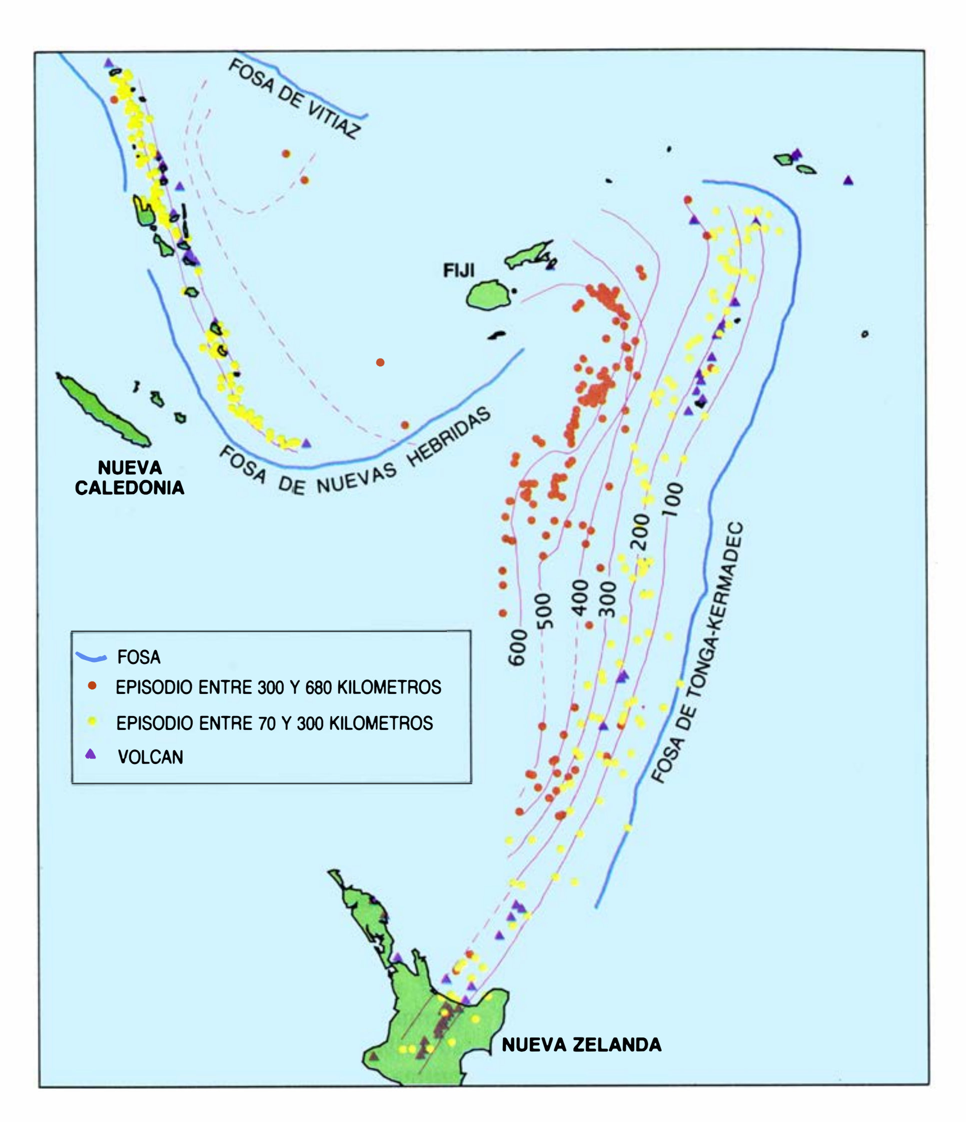
6. ESTADISTICA DE LOS TERREMOTOS en función de la profundidad. El número de terremotos de magnitud 5 o mayor en cada intervalo de 10 kilómetros de profundidad (*diagrama de la izquierda*) alcanza un mínimo a unos 400 kilómetros, pero luego vuelve a aumentar, antes de caer abruptamente a cero a unos 650 kilómetros. Las réplicas son raras en la mayoría de los episodios profundos de tamaño moderado, pero se vuelven más comunes a las máximas profundidades (*diagrama del centro*). Los terremotos que desarrollan sacudidas más fuertes suelen registrarse a profundidades someras; a las mayores, la magnitud de los episodios más fuertes se mantiene muy constante hasta que la actividad sísmica se anula del todo (*diagrama de la derecha*). Los datos sugieren que los cambios en la fase cristalina de rocas del manto, que se cree que se producen a profundidades de 400 y 650 kilómetros, pueden afectar a los terremotos profundos.

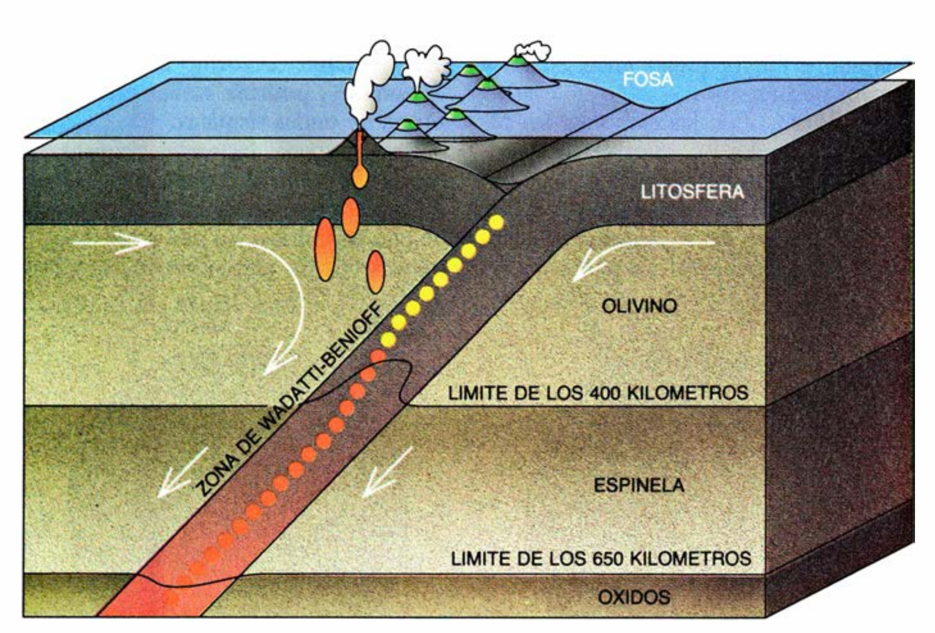
brinda varias posibilidades en cuanto a las fuentes del esfuerzo que se libera en terremotos profundos. Una placa descendente experimenta deformaciones variadas, se dobla, se estira, se comprime; el calentamiento y los cambios de fase también podrían generar esfuerzos, al cambiar el volumen de la roca. ¿Cómo se liberan esos esfuerzos? ¿Qué ocurre realmente en el foco de un terremoto profundo?

La magnitud de los terremotos profundos puede equipararse a la de cualquier otro, si exceptuamos los someros muy grandes. El terremoto profundo de mayor potencia destructora registrado en los últimos años, el episodio rumano de 1977, tuvo una magnitud de 7,2, y otro de magnitud 7,6 se produjo a 650 kilómetros bajo Colombia, en 1970. Sin embargo, el mecanismo de liberación de energía no debe ser idéntico a la ruptura frágil que desencadena los terremotos someros. Aun cuando el material en el que ocurren los terremotos profundos esté mucho más frío —y, por tanto, más rígido— de lo que Jeffreys creyó, no debería fracturarse según las pautas seguidas por las rocas a baja presión. Suponiendo que el esfuerzo abriera una grieta, el peso de toda la roca situada encima de ella la volvería a sellar. Si la roca se deforma, lo hará plásticamente. Se mantiene la objeción de Jeffreys: los terremotos convencionales, en los que la roca se rompe y se desliza, no deberían ocurrir en el manto.

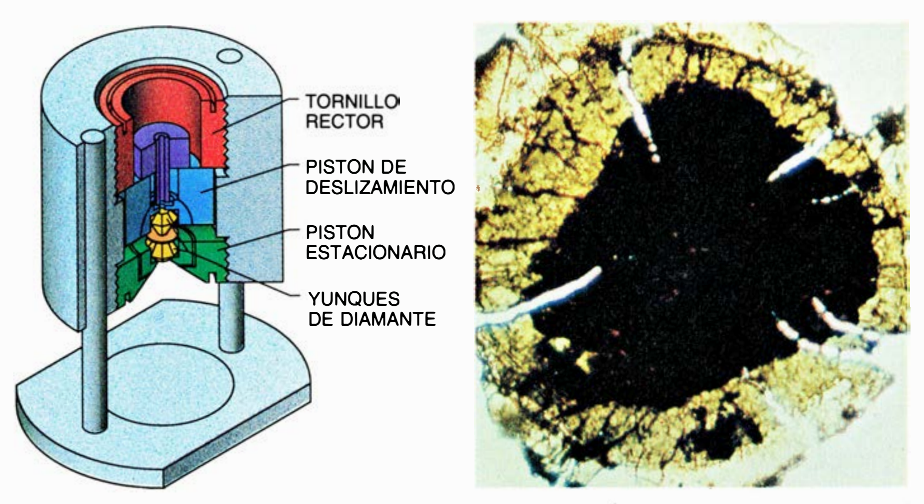
Estudios sismográficos recientes de terremotos profundos apoyan la argumentación que rechaza el mecanismo convencional. Casi todos los grandes episodios someros se ven acompañados de muchos temblores menores. Estas réplicas, así se llaman, suelen ocurrir a lo largo del mismo plano que el deslizamiento inicial, liberando aparentemente tensiones residuales en la línea de la fractura. En los terremotos profundos, las réplicas escasean. El terremoto profundo de 1970, bajo Colombia, el mayor quizá de los episodios realmente profundos de los últimos 25 años, no presentó ningún tipo de réplicas. Los terremotos profundos que tienen réplicas —y éstas son más comunes entre los episodios más profundos— poseen, por lo general, unas cuantas como máximo.

Las réplicas que se producen definen un patrón espacial completamente diferente del que dan los terremotos someros. Recientemente, Raymond J. Willermann, entonces en el Laboratorio Nacional de Los Alamos, y yo estudiamos las relaciones espaciales entre





7. EN LA ZONA DE SUBDUCCION se localizan casi todos los terremotos profundos. La hondura focal de los registrados a lo largo de la fosa de Tonga-Kermadec, una fosa oceánica del Pacífico sudoccidental, bordeada por actividad sísmica e islas volcánicas, se sitúa a lo largo de una serie de curvas de nivel paralelas de profundidad creciente: patrón conocido como zona de Wadati-Benioff (*mapa superior*). La zona de Wadati-Benioff señala la subducción de una placa litosférica (*abajo*): los terremotos se producen en la laja descendente. El doblamiento hacia abajo de la placa subducente es responsable de la fosa: el material fundido que asciende desde la placa alimenta la alineación de volcanes. El proceso se halla gobernado por la circulación convectiva del manto; la placa descendente constituye la rama fría, que se va sumergiendo, del circuito.



8. CAMBIOS DE FASE EN LAS ROCAS DEL MANTO. Podrían desempeñar cierta función en los terremotos profundos. Se simulan en una celda de yunque de diamante, aparato que comprime muestras de roca entre dos diamantes (*izquierda*). La transparencia de los diamantes permite calentar las muestras con un láser y fotografiarlas. En una muestra de olivino (constituyente principal del manto), comprimida hasta 300.000 atmósferas y calentada a 1500 grados Celsius, las fases diferentes forman anillos concéntricos (*derecha*). Un anillo externo pálido de olivino inalterado da paso a una espinela más amarillenta, en una transición que, según se cree, ocurre a una profundidad de 400 kilómetros; en el centro, donde la presión y la temperatura son máximas, se halla la oscura fase de óxido, a la que debe cambiar la espinela a 650 kilómetros. La fotografía de la derecha ha sido cedida por William A. Bassett, adscrito a la Universidad de Cornell.

las sacudidas iniciales y las réplicas de terremotos profundos. Los terremotos someros pequeños suelen tener réplicas centradas relativamente cerca del evento principal, hecho que concuerda con la idea de que las réplicas representan un deslizamiento continuado a lo largo de la misma fractura que produjo el temblor principal. Encontramos, sin embargo, que algunos terremotos profundos pequeños —los de magnitud 5,5 o menor— tienen réplicas a una distancia de 30 kilómetros, o más, de la sacudida inicial. No es probable que una zona de ruptura responsable de un terremoto tan pequeño mida 30 kilómetros de longitud.

A diferencia de las someras, las réplicas profundas no caen en la orientación de un plano. Los datos existentes, por el contrario, sugieren una distribución más o menos aleatoria en torno al episodio inicial. La pauta observada revela, de nuevo, que los terremotos profundos y los someros tienen mecanismos fundamentalmente diferentes.

Una forma atractiva, aunque incorrecta, de explicar los terremotos profundos es la que se remonta casi a su fecha de descubrimiento. Según la misma, son resultado directo de la transformación de material subducido en una fase más densa. Tales transformaciones debían tener lugar en la roca subducida y, si se producían con la rapidez suficiente —si la roca en efecto implosionaba conforme aumentase su densidad—, podrían radiar energía en forma de ondas sísmicas.

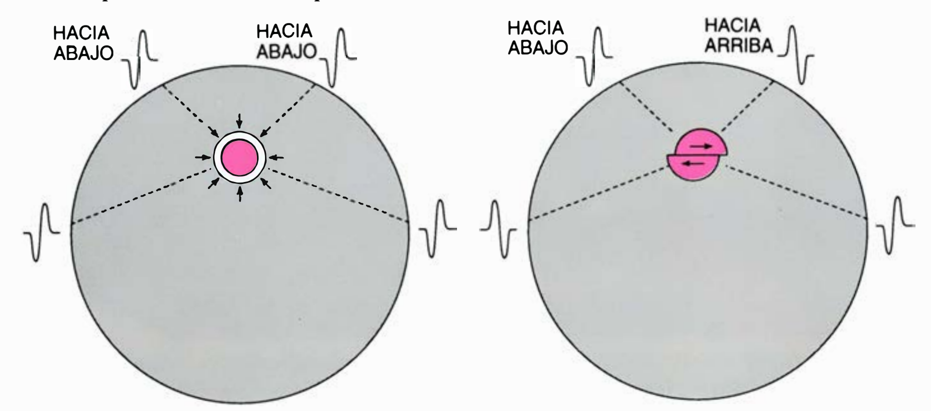
Por desgracia, las ondas sísmicas detectadas a partir de terremotos profundos no se parecen en nada a la rúbrica de una implosión. En una implosión, el material de los alrededores se desplaza hacia el interior, hacia el foco. Cabría esperar, pues, que todos los sismogramas (que, con la amplitud, detectan también la dirección de las ondas sísmicas) registraran un movimiento inicial hacia abajo, hacia el interior de la Tierra durante el episodio. Además, dado que una implosión produce movimiento radial, más que transversal, generaría ondas *P* mucho más marcadas que las *S*.

Los movimientos primeros de un terremoto profundo se dirigen hacia abajo en algunas áreas y hacia arriba en otras, lo mismo que en los terremotos someros. Los movimientos hacia arriba y hacia abajo divergen: diríase que una parte de la Tierra se hubiera movido en una dirección, a lo largo de un plano de deslizamiento, y la otra parte se hubiera desplazado en sentido opuesto; igual pauta de comportamiento observamos en sismogramas de episodios someros. En los terremotos profundos, lo mismo que en los someros, las ondas *S* son, también, mucho más fuertes que las ondas *P*; ello sugiere que la fuente no es una implosión, sino un deslizamiento.

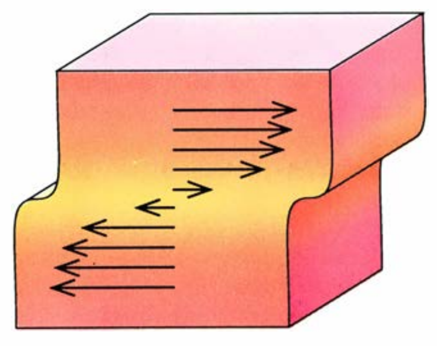
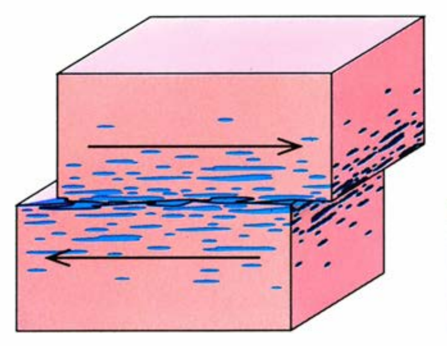
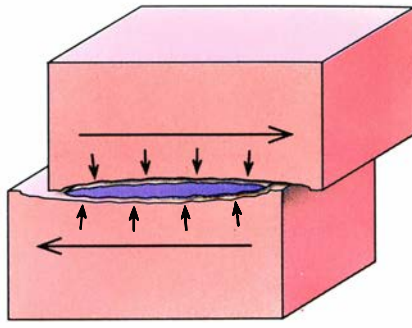
Pero, ¿cómo puede la roca deslizarse bruscamente, si las enormes presiones del manto descartan la ruptura frágil? Un esquema propuesto, en los años sesenta, por David T. Griggs, de la Universidad de California en Los Angeles, y elaborado recientemente por Masaki Ogawa, de la Universidad de Tokyo, postula desechar la deformación dúctil. Un terremoto profundo podría producirse cuando la roca que se deforma bajo esfuerzo de cizalla empieza a generar calor de rozamiento a un ritmo mayor que el que la roca circundante puede disipar. El calor reblandece la roca e incluso funde parte de la misma, acelerando la deformación. El proceso de realimentación podría provocar que la temperatura y la tasa de deslizamiento aumentaran explosivamente y produjeran un terremoto.

La verosimilitud del mecanismo depende mucho de la composición y estructura de la roca en una zona de Wadati-Benioff. Se privilegia cuando la estructura de la roca —verbigracia, una capa débil preexistente— tiende a concentrar el deslizameiento dúctil a lo largo de un plano. No es cierto en absoluto que la disposición estratiforme del material subducido tenga la orientación adecuada para favorecer el deslizamiento en los sentidos más frecuentemente observados en los terremotos profundos.

Otro cuadro de explicación diferente atribuye el deslizamiento en zonas de terremotos profundos al efecto de in-



9. ENTRE LAS CAUSAS de los terremotos profundos debe descartarse la implosión debida a la súbita transformación de roca subducida en una fase más densa. La implosión tendría que manifestarse en los sismógrafos de todo el mundo como un movimiento inicial hacia abajo (*izquierda*). En cambio, los terremotos profundos generan un movimiento inicial hacia arriba en algunos puntos y hacia abajo en otros (*derecha*), lo que induce a pensar que su fuente es el deslizamiento lateral en rocas de gran profundidad.

10. EN LOS MECANISMOS DE DESLIZAMIENTO PROPUESTOS subyace el reconocimiento de que, a grandes profundidades, la Tierra no puede ceder por simple fractura. Según determinada hipótesis, la lenta deformación de la roca se acelera súbitamente con la acumulación de calor de rozamiento, lo que reblandece la roca y aumenta la velocidad de deformación en un proceso sin freno (*izquierda*). Un segundo mecanismo atribuye el deslizamiento a la influencia de fluidos (*centro*). Por debajo de cierta profundidad, la elevada presión debe liberar agua aprisionada en la estructura cristalina de minerales subducidos; el agua debe contrarrestar fuerzas que traban fallas potenciales, permitiéndoles ceder. La tercera propuesta sostiene que los esfuerzos de cizalla inducen un cambio de fase en una capa paralela al esfuerzo (*derecha*). El súbito cambio de la estructura cristalina debilitaría la roca, permitiendo el deslizamiento.

clusiones fluidas. Experimentos de laboratorio han demostrado que, a presiones equivalentes a las de las profundidades más someras, los fluidos aprisionados en los poros de la roca pueden contrarrestar las fuerzas que traban una fractura potencial, lo que permite que ceda aquélla ante un esfuerzo de cizalla menor. Por lo menos en un caso, en el Arsenal de las Montañas Rocosas, cercano a Denver, se produjo una sucesión de terremotos someros tras la inyección de fluidos residuales en el terreno, lo que redujo, según parece, los esfuerzos de confinamiento y las capas de roca se deslizaron.

En 1966, C. B. Raleigh y Mervyn S. Paterson, de la Universidad Nacional Australiana, sugirieron que los fluidos de los poros en zonas de terremotos profundos deberían producir el mismo efecto, permitiendo que el material subducido se rompiera como lo hacen las rocas a presiones inferiores. La fuente de los fluidos, propusieron Raleigh y Paterson, debía ser la deshidratación de la serpentina (una forma de silicato magnésico) u otros minerales parecidos en el material subducido: la liberación del agua incorporada en su estructura cristalina, al calentarse por contacto con el manto circundante, a más de 500 grados Celsius. Raleigh y Paterson propusieron también otras fuentes de fluido: el agua atrapada en los sedimentos de las fosas marítimas profundas y arrastrada hacia abajo con la corteza, amén de la roca del manto parcialmente fundida.

Para condicionar las propiedades globales de la masa rocosa, piénsese en su tendencia a la fractura, un fluido ha de tener la posibilidad de migrar por su interior; y no es seguro que la roca del manto sea suficientemente porosa. Además, cabe esperar la deshidratación de minerales a temperaturas y presiones específicas, correspondientes a profundidades también específicas. Si los fluidos de los poros desencadenasen terremotos profundos, esas regiones deberían mostrar intensas concentraciones de sismicidad. Sin embargo, la frecuencia de los terremotos muestra que la variación con la profundidad es sólo moderada.

La variación que hay (la anulación de la actividad sísmica en el límite olivino-espinela, a 400 kilómetros, y su reanudación a profundidades mayores) parece guardar cierta relación con las profundidades de cambios de fase. Stephen H. Kirby, del Servicio Geológico de Estados Unidos, ha propuesto un mecanismo de terremoto profundo que depende de transiciones de fase, aunque, a diferencia de modelos anteriores, produce deslizamiento en vez de implosiones. A modo de sustitutos de la roca real del manto, Kirby y sus colaboradores estudiaron hielo y tremolita, un silicato de calcio y magnesio; ambos cambian a una fase más densa a presiones asequibles en el laboratorio.

Cuando los investigadores comprimieron cada material a una presión levemente inferior a la normal de transición de fase, y lo sometieron a esfuerzo, encontraron que la transición de fase se había desencadenado a lo largo de una delgada capa paralela a la tensión. La repentina reordenación de la estructura cristalina a lo largo de la capa debilitó aparentemente el material, permitiéndole deslizarse. Kirby y sus colegas notaron que, en tal proceso, sus muestras emitían ruidos o chasquidos de rotura, equivalentes, en el laboratorio, a los de terremotos.

Kirby sostiene que estos cambios de fase prematuros acontecen también en roca subducida, como si ésta se hallara sometida a esfuerzos; afirma, asimismo, que el deslizamiento resultante explica, por lo menos, algunos terremotos en las zonas de Wadati-Benioff. Esos puntos de vista no se oponen a la existencia de terremotos en una amplia gama de profundidades superiores a la del límite de 400 kilómetros: diversos investigadores, incluido William A. Bassett, de la Universidad de Cornell, han encontrado que las transiciones de fase en material subducido pueden ocurrir a diversas profundidades, en función de la composición precisa de aquél y de su velocidad de descenso. Con todo, el mecanismo de Kirby explicaría la súbita desaparición de terremotos profundos más allá de los 680 kilómetros. A esas honduras se han desarrollado ya todas las transiciones de fase del manto conocidas.

Nadie ha demostrado todavía que el esfuerzo de cizalla produzca el mismo efecto en las transiciones de fase de roca real del manto que el que provoca en hielo y tremolita. Mas, aun cuando la hipótesis de Kirby fuera errónea, las transiciones de fase pueden desempeñar cierta función en los terremotos profundos. A lo mejor generan sólo tensiones que se liberan súbitamente en otra parte, donde las rocas ceden sin que sepamos cómo.

Tal vez pronto se pueda elegir de entre estas hipótesis con un mayor margen de confianza. Raymond Jeanloz y Charles Meade, de la Universidad de California en Berkeley, están recreando condiciones de manto en el laboratorio, para estudiar los mecanismos de terremoto profundo propuestos. Una prensa del tamaño de un puño, llamada celda de yunque de diamante, genera las presiones necesarias estrujando una minúscula muestra de roca entre las puntas de dos diamantes. La muestra puede calentarse enviando un haz láser a través de uno de los diamantes; las transiciones de fase y otros cambios de la roca pueden verse a través del otro. Mientras tanto, los sensores acústicos detectan cualquier "terremoto". Las

investigaciones, en sus etapas iniciales, sugieren ya que, a altas presiones, el olivino sólo cede cuando además contiene serpentina; este resultado corrobora la hipótesis del mecanismo de deshidratación postulado por Raleigh y Paterson.

Demostrando, respectivamente, la realidad de los terremotos profundos y su "imposibilidad", Wadati y Jeffreys plantearon un rompecabezas cuya resolución mantiene a los geofísicos en vilo. (Y siguen todavía vigentes, transcurridos más de sesenta años de su muerte.) En el contexto de la tectónica de placas y la convección del manto, que los propios terremotos profundos ayudaron a establecer, estos acontecimientos han planteado nuevos interrogantes.

Primero, la desaparición de terremotos a 680 kilómetros de profundidad. ¿Marca esa cifra el límite de convección del manto? ¿Se trata acaso de un mero cambio en las propiedades mecánicas de un manto que está en todo su grosor sujeto a convección? Los terremotos profundos ocasionales ocurridos en regiones desprovistas de zonas de subducción conocidas plantean su propio enigma. Los terremotos profundos de Rumanía y del Hindu Kush, dos de esas regiones, pueden reflejar la presencia de una antigua zona de subducción enmascarada por actividad tectónica posterior. Esta explicación es menos plausible para los temblores registrados bajo Africa septentrional y España. Aquí, el misterio de los terremotos profundos está envuelto con otro velo de oscuridad: la posibilidad de que algunos terremotos profundos tengan lugar sin subducción ninguna.

BIBLIOGRAFIA COMPLEMENTARIA

SHALLOW AND DEEP EARTHQUAKES. K. Wadati en *The Geophysical Magazine*, vol. 1, n.º 4, págs. 162-202; marzo de 1928.

DEEP-FOCUS EARTHQUAKES AND THEIR GEOLOGICAL SIGNIFICANCE. Andrew Leith y J. A. Sharpe en *The Journal of Geology*, vol. 44. n.º 8, págs. 877-917; noviembre-diciembre de 1936.

KIYOO WADATI AND EARLY RESEARCH ON DEEP FOCUS EARTHQUAKES: INTRODUCTION TO SPECIAL SECTION ON DEEP AND INTERMEDIATE FOCUS EARTHQUAKES. Cliff Frohlich en *Journal of Geophysical Research*, vol. 92, n.º B13, págs. 13777-13788; 10 de diciembre de 1987.

LOCALIZED POLYMORPHIC PHASE TRANSFORMATIONS IN HIGH-PRESSURE FAULTS AND APPLICATIONS TO THE PHYSICAL MECHANISM OF DEEP EARTHQUAKES. Stephen H. Kirby en *Journal of Geophysical Research*, vol. 92, n.º B13, págs. 13789-13800; 10 de diciembre de 1987.

Ciencia y sociedad

Archaeopterix sometido a TAC

Se han recuperado hasta la fecha seis restos esqueléticos y una pluma de *Archaeopteryx*. Todos ellos proceden de los yacimientos de calizas litográficas de Solhofen, en el mediodía del Jura bávaro. Con una edad cercana a los 150 millones de años, constituyen las aves más antiguas del mundo. De ahí el interés que *Archaeopteryx* ha despertado entre paleontólogos y teóricos de la evolución; el animal conjuga caracteres de reptil y de ave que permiten esbozar la filogenia de las aves en términos generales.

Las posibilidades de análisis de los esqueletos fósiles dependen de la calidad de su conservación, así como de la preparación idónea de los mismos. Conviene desprender, con sumo cuidado, los elementos esqueléticos fósiles de la matriz pétrea que los abriga. Se hace de manera manual, mecánica, aplicando chorros de arena o mediante tratamiento químico con ácidos. No cabe, sin embargo, la extracción completa de los huesos en los fósiles de Solnhofen, pues se hallan incrustados y embebidos: la sustancia ósea, transformada en calcita, se desharía en migajas al primer intento de separación. A pesar de todo, se ha buscado cómo conseguir una imagen espacial (igual que en el ejemplar *Archaeopteryx* de Londres), e incluso limpiar el fondo de la matriz rocosa, sin dejar de lado otras técnicas de análisis: luz ultravioleta, infrarroja y de rayos X.

Para aclarar las cuestiones sistemáticas, de parentesco y filogenia, el cráneo del vertebrado reviste un interés primordial, ya que sus rasgos estructurales permiten distintos análisis. Pero no todos los ejemplares de *Archaeopteryx* conservan el cráneo. Sólo los de Berlín y Eichstätt lo poseen entero.

Sin embargo, la sustancia ósea ha sufrido, también aquí, un aplastamiento total durante el proceso de petrificación, de modo que varios detalles craneanos (especialmente en la región posterior) no han podido recibir, hasta ahora, la adecuada descripción morfológica. Las imágenes de rayos X aportaron poca información, ya que las imágenes que reproducen se basan en la perpendicular del rayo a través del fósil.

Así las cosas, aparece la tomografía axial computerizada. Ideada en los años setenta para conseguir imágenes axiales de las capas craneales humanas, se ha venido perfeccionando para su aplicación médica y método de investigación del cuerpo. Las imágenes se obtienen mediante un sistema rotatorio de precisión de un tubo de rayos X y varios detectores colocados al otro lado del tubo, frente al plano de rotación. El debilitamiento sufrido por el paso de los rayos X a través del objeto, registrado por los detectores, llega a un ordenador. Este divide la imagen del objeto, obtenida por el paso de los rayos, en una matriz de elementos de pequeño volumen y evalúa, para cada uno de ellos, la desviación relativa de sus coeficientes lineales de debilitamiento con relación al valor correspondiente del agua, conocido como valor-CT. Estos valores se transforman en distintos tonos de gris, que servirán para crear la imagen de la sección.

La tomografía axial computerizada llegó a la paleontología en 1984; se aplicó para medir el volumen de cavidades endocraneales de primates, fuertemente solidificados. Esto sólo es posible cuando se encuentran cráneos sin deformar (fenómeno insólito en paleontología). A pesar de las deformaciones del cráneo de *Archaeopteryx* de Eichstätt, parece, sin embargo, que existe una posibilidad real de obtener información, mediante la técnica de reconstrucción de corte secundario realizada con tomografía axial computerizada, sobre ciertas estructuras del interior de la cavidad craneal.

Abordamos el estudio del *Archaeopteryx* de Eichstätt con un tomógrafo axial computerizado Siemens Somaton DR, versión H, en colaboración con Mathias Prokop y Wilfried Döhring, de la Facultad de Medicina de Hannover, y John H. Ostrom, de la Universidad de Yale en New Haven (Connecticut) en los laboratorios de Siemens AG en Erlangen.

Se sacaron imágenes del cráneo y de la matriz de roca circundante en cuarenta y seis cortes continuos, paralelos entre sí, de un milímetro de espesor, para así observar los huesos de la mandíbula inferior. El proceso de barrido duró siete segundos; la tensión en el tubo alcanzó ciento veinticinco kilovolt. El tamaño de los elementos de volumen diferenciables correspondió a un cubo de $0{,}11 \times 0{,}11 \times 1{,}0$ milímetros. La utilización de una matriz de imagen de 512×512 permitió una resolución de imagen cercana a 0,5 milímetros. Sin embargo, la aparición de artefactos, rayas y otras interferencias, generados por las grandes diferencias de

Cráneo del Archaeopteryx *de Eichstätt, procedente de las calizas litográficas del Jurásico Superior en Workerszell. El hueso cuadrado derecho queda visible en el lado posterior del cráneo, entre la caja encefálica y la articulación de la mandíbula inferior.*

densidad entre las placas de roca y el aire circundante, hizo imposible la interpretación directa de la imagen obtenida. Estas interferencias en las imágenes pudieron eliminarse mediante la obtención de reconstrucciones de corte secundario. El conjunto de datos de los elementos de volumen almacenados posibilitó reconstrucciones de las imágenes paralelas a la superficie de la placa o en ángulos arbitrarios con respecto a la misma.

Estas imágenes oblicuas abrieron nuevas perspectivas frente a las técnicas habituales de rayos X. Lo que sirvió, sobre todo, para conocer la forma del hueso cuadrado, que en las aves modernas está unido, en su parte superior, mediante dos cóndilos articulares, a la caja encefálica y, en su parte inferior, se produce la unión articular con la mandíbula inferior. Los reptiles tienen un hueso cuadrado con un único cóndilo articular superior.

Hasta ahora, se había supuesto también que, en *Archaeopteryx*, dicho hueso era monocondilar y, por tanto, evolutivamente de tipo reptiliano. En el ejemplar de Eichstätt se observó sólo un cóndilo articular superior, no sólo mediante la inspección microscópica, sino también mediante una inspección por rayos X.

El método de tomografía axial computerizada permite colocar un objeto y obtener imágenes bidimensionales en la dirección deseada, sin interferencias, a partir de una matriz de datos tridimensionales; las representaciones tienen 0,2 milímetros de grosor. Las imágenes, en un ángulo de treinta y dos grados con respecto a la superficie del fósil, muestran que la porción superior del hueso cuadrado derecho está dotada en realidad de dos cóndilos articulares.

Aparece así un nuevo carácter exclusivo de las aves en *Archaeopteryx*, además de las plumas y la fúrcula. Este descubrimiento coincide con la conclusión, a la que llegan también otros investigadores, según la cual la cápsula auditiva de *Archaeopteryx* muestra una estructura típica de las aves. Dadas las estrechas relaciones topográficas y funcionales del hueso cuadrado con los elementos del cráneo que forman la cápsula auditiva, no es de extrañar que en el *Archaeopteryx* ambas estructuras queden trabadas igual que en las aves.

Nuestra interpretación de la articulación del hueso cuadrado de *Archaeopteryx* y, en especial, la posibilidad de identificar qué huesos de la caja encefálica y de la cápsula auditiva se hallan unidos a los cóndilos articulares del hueso cuadrado, se mueve en un terreno todavía poco seguro, incluso con la resolución que obtenemos mediante la realización técnica de la tomografía axial computerizada. Sin embargo, queda fuera de toda duda el valor potencial de esta técnica para investigar la morfología del cráneo y de otros fósiles que sean ejemplares únicos. (Bernd Haubitz y Peter Wellnhofer.)

La fusión nuclear a debate

Corren buenos tiempos para la investigación en el campo de la energía de fusión ante las alteraciones del clima provocadas, probablemente, por el continuo vertido de gases a la atmósfera, consecuencia del consumo de combustibles fósiles. Preocupación a la que debemos sumar la relativa al tratamiento y almacenamiento de los residuos nucleares de fisión. En un orden complementario, los investigadores se muestran esperanzados en la posibilidad de generar, en un reactor de fusión, más energía de la que éste consume. A pesar de todo, hasta los más optimistas admiten las graves dificultades que se esconden tras ese espejismo de bonanza: queda por resolver uno de los principales interrogantes, a saber, la complejidad –y coste consiguiente– de los reactores de fusión.

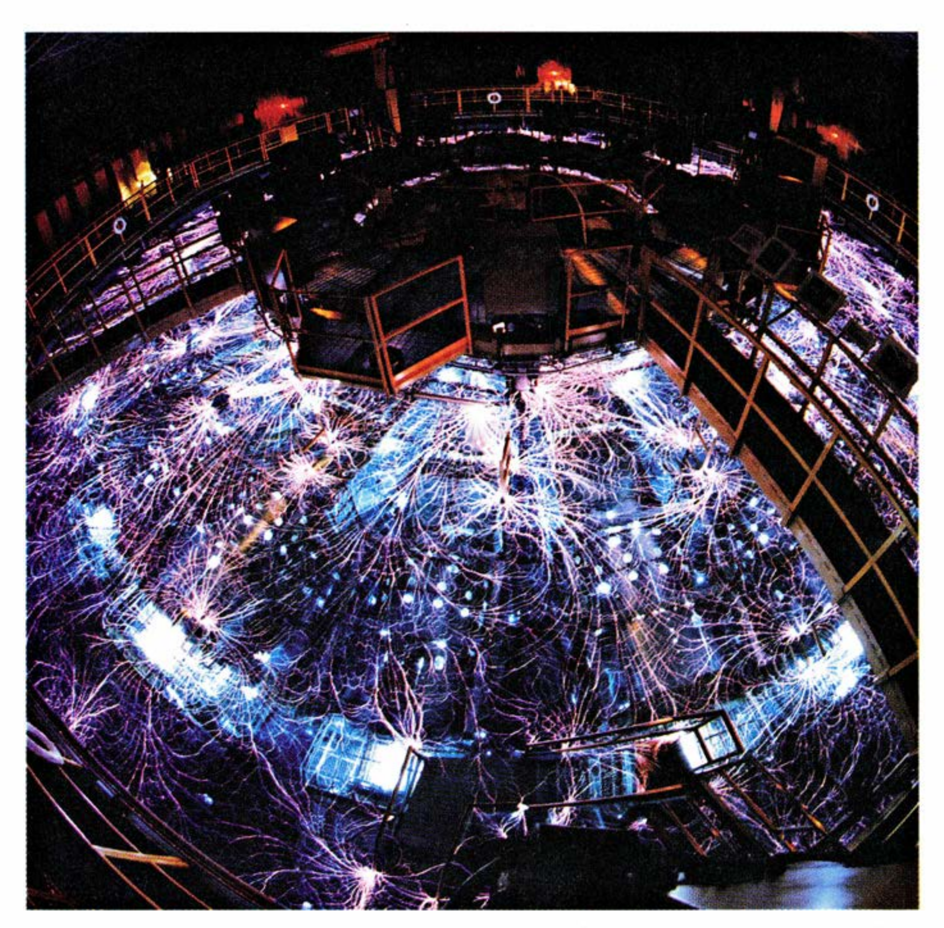
Años atrás, imaginábanse los reactores de fusión máquinas seguras que consumirían un combustible abundante y barato. El cuadro ha cambiado sustancialmente con el tiempo gracias al mejor conocimiento del desarrollo de las centrales nucleares de fusión. Necesitarán un combustible caro y emitirán radiactividad. El tamaño y la complejidad de una central nuclear de fusión serán parecidos, si no mayores que los de nuestros generadores de fisión. Los gobiernos, a su vez, se mostrarán reacios a invertir miles de millones de dólares en la construcción de generadores gigantescos; preferirán apostar, a largo plazo, por técnicas de riesgo y escala menores, aun cuando funcionaran los generadores previstos. Cierto que las principales potencias industriales, los EE. UU., la Unión Soviética, el Japón y algunos países de la Comunidad Europea siguen favoreciendo la investigación en ese campo; pero no lo es menos que los presupuestos menguan en cada año fiscal.

Los físicos pretenden "domesticar" lo que hace lucir al sol y explotar las bombas termonucleares. Mientras que la fisión nuclear consiste en la rotura de núcleos de átomos pesados, el proceso de fusión nuclear se da cuando dos núcleos de átomos ligeros, hidrógeno o helio por ejemplo, se unen para formar un núcleo más pesado, liberando energía. Si bien se han propuesto, durante los últimos 40 años, numerosos métodos para producir la fusión nuclear, sólo sobreviven dos: el confinamiento magnético y el confinamiento inercial.

Se empezó a trabajar en el confinamiento magnético en los primeros años de la década de los cincuenta; hoy constituye el procedimiento seguido en más de dos tercios de las investigaciones. Consiste en calentar y comprimir un plasma (un gas ionizado que contiene proporciones parecidas de electrones y de iones), confinado por un campo magnético. A finales de los sesenta, los físicos soviéticos desarrollaron un recinto en forma toroidal, el tokamak, que supuso un importante paso hacia adelante. En el tokamak de Princeton y en el de Culham, aquél en Estados Unidos y éste en Inglaterra, se han creado plasmas cuya temperatura supera la reinante en el interior solar, aunque sólo por un instante; en efecto, el plasma se inestabiliza enseguida y se dispersa, relajando su temperatura y densidad y contaminándose en su contacto con el recinto que lo aloja. Estas máquinas se hallan a un orden de magnitud del punto de ignición, es decir, de aquella situación en que la reacción de fusión produce calor suficiente para que la reacción de mantenga sin necesidad de recabarlo fuera de la máquina.

Hasta ahora, los plasmas han sido de deuterio, isótopo pesado del hidrógeno, abundante en el agua de mar y no radiactivo. Se creyó que el deuterio constituiría el combustible exclusivo, pero se ha comprobado que la fusión del deuterio no produce energía suficiente. Pero empieza a admitirse que valdría más un combustible formado por deuterio y tritio, isótopo éste del hidrógeno y promotor de la explosión de la bomba termonuclear. La fusión de una mezcla de deuterio y tritio produce 200 veces más energía que la fusión entre átomos de deuterio. No importa la escasez del tritio en la naturaleza; los reactores nodriza que bombardean litio con neutrones rápidos pueden fabricarlo. Puesto que la reacción de fusión del deuterio con el tritio produce neutrones rápidos, un reactor de fusión revestido de litio podría generar su propio tritio; el calor producido en esta reacción podría también destinarse a la creación del vapor que moviera las turbinas.

Pero los pasos que hemos dibujado no están libres de inconvenientes. La producción del combustible no es ninguna técnica sencilla, y prueba de ello nos la ofrecen las dificultades con que se encuentran los reactores destinados



LAS DESCARGAS ELECTRICAS se entrecruzan en la superficie del baño de agua del Acelerador del Laboratorio Nacional de Sandia en la producción de pulsos de iones ligeros. Una máquina de estas características serviría de elemento cebador de un sistema de confinamiento inercial avanzado. Sus haces de un millón de joule superan en potencia a cualquier láser.

a este menester para uso militar. Ni sería tarea fácil proyectar un reactor con una cubierta de litio, teniendo en cuenta el alto poder reactivo de este elemento que lo convierte en explosivo. Además, el tritio es radiactivo y los neutrones rápidos de la reacción deuterio-tritio contaminarían de radiactividad todo el recinto. Aunque esa radiactividad no alcanzase el poder letal, ni la persistencia de la producida por los reactores de fisión actuales, el conjunto de medidas necesarias para su contención, tanto durante la vida operativa del reactor como después, encarecerían el coste total de la máquina.

Algunos investigadores abogan por otro combustible: la mezcla de deuterio con helio 3, isótopo no radiactivo del helio. La reacción de fusión tiene un alto rendimiento y es limpia, es decir, no produce neutrones. El helio 3, escaso en la tierra, puede fabricarse de elementos ligeros y neutrones rápidos. Y dada la gran abundancia de helio 3 en la Luna, donde el viento solar deposita ingentes cantidades del mismo, se ha propuesto también el beneficio de dicho elemento en el satélite. El problema que aquí se plantea, sin embargo, reside en la temperatura: la ignición de un reactor con deuterio y helio 3 requiere una temperatura mayor y, por tanto, un recinto más potente y caro que en el caso de emplear una mezcla de deuterio-tritio.

Existe un proyecto oficial para la construcción, en Princeton, de un reactor de deuterio-tritio. El coste del proyecto, que entrará en función en 1996, se cifra en 400 millones de dólares. La verdad es que esa máquina no está llamada a ser ni siquiera un prototipo, pues no podrá transformar la energía de fusión en electricidad; sólo podrá alcanzar la ignición durante cinco o diez segundos a lo sumo. Estados Unidos, la Unión Soviética, Japón y Europa se hallan estudiando la posibilidad de construir un auténtico prototipo de generador, que empezaría a operar a principios del siglo XXI. De acuerdo con el programa previsto, en el próximo decenio un equipo internacional diseñará el prototipo y los países participantes decidirán entonces la oportunidad de su construcción.

Lo que sí han comenzado ya son las discrepancias sobre aspectos básicos del proyecto. Muchos miembros del equipo son partidarios de construir un reactor de altas prestaciones, con imanes superconductores, que generaría mil millones de watt de energía térmica, el rendimiento del mayor reactor de fisión operativo. Por encima de esa capacidad, estiman los expertos, la máquina sobrepasaría las posibilidades de los gobiernos y los complejos industriales. No comparte ese criterio Paul H. Rebut, director ingeniero del Toro Conjunto Europeo, considerado el tokamak más avanzado del mundo; él se muestra partidario de construir un reactor capaz de producir tres mil millones de watt utilizando imanes convencionales. En su opinión, un reactor de menores dimensiones tal vez no funcionara. Para ciertos observadores, tampoco está claro que participen los Estados Unidos en la construcción del reactor internacional; por dos razones: su elevado coste (2000 millones de dólares apuntando bajo) y su temor ante la posibilidad de transferencia de técnicas avanzadas a otros países, la Unión Soviética en particular.

El campo del confinamiento inercial entró en escena a mediados de los años sesenta, relacionado con la investigación en láseres. Hoy, se encuentra inmerso en disputas técnicas y políticas. Si el confinamiento magnético busca imitar la conflagración en el interior de una estrella, el confinamiento inercial pretende crear una explosión termonuclear en miniatura. Se trata de detonar una pequeña cápsula esférica de deuterio y tritio con radiación. La implosión resultante calienta y aumenta tanto la densidad del combustible que provocan su ignición. El objetivo oficial del programa de confinamiento inercial no es producir una nueva fuente de energía, sino entender mejor la tecnología nuclear de aplicación militar. De ahí que consideren reservados muchos aspectos y, en Estados Unidos por ejemplo, esté prohibida cualquier colaboración internacional al respecto. Pero muchos investigadores en confinamiento inercial defienden que lo único que les guía es producir "energía para el planeta".

Se creyó en un principio que bastaría con la energía de una luz láser de unos centenares de joule para producir la ignición (una bombilla de 100 watt produce 100 joule en un segundo). Con el paso del tiempo las estimaciones aumentaron, junto con la desconfianza en las posibilidades de esta técnica. Hasta que, recientemente, los resultados de experimentos secretos realizados en Nevada, bajo el nombre cifrado de

Centurion-Halite, han revitalizado este campo de trabajo. Se ha producido la ignición de cápsulas de deuterio-tritio a partir de los rayos X procedentes de explosiones nucleares subterráneas. Los partidarios del confinamiento inercial afirman que estos experimentos aportan la prueba esencial que no se ha podido conseguir en las experiencias de confinamiento magnético. Todo lo que hay que hacer ahora, según ellos, es construir un elemento cebador, una fuente de radiación, capaz de producir los mismos resultados. A este propósito, se comienza a hablar ya de construir un laboratorio de microfusión con ese objetivo.

Diversos laboratorios han entrado en liza por la creación de ese elemento cebador, aunque no se ha estudiado todavía ni la potencia ni la forma que debe tener. Para la mayoría de los expertos, las pruebas Centurion-Halite revelan que la ignición de los blancos no precisa más allá de 10 millones de joule de energía. Esta cifra centuplica la energía emitida por el mayor láser construido hasta ahora, el láser de vidrio de neodimio, Nova, del laboratorio nacional Livermore. Según Erik Storm, supervisor del proyecto de confinamiento inercial en este laboratorio, una versión mayor del Nova sería el mejor candidato para producir la microfusión, ya que ha sido este láser el instrumento que ha conseguido mayores éxitos en la compresión y en el calentamiento de blancos.

Opinión de la que se apartan otros físicos, que sostienen que esa clase de láseres resulta ineficaz (el Nova, por ejemplo, no llega a convertir en energía luminosa ni un 1 por ciento de la energía eléctrica que consume) y proclive a sobrecalentamientos. Sin olvidar su carestía: el coste del Nova superó los 200 millones de dólares, lo que induce a suponer que otro láser 100 veces más potente decuplicaría fácilmente su coste. Se ha especulado sobre otros sistemas para producir la microfusión: láser de gas de criptón-flúor, del Laboratorio Nacional de Los Alamos, el acelerador de iones ligeros, del Laboratorio Nacional de Sandia, y el acelerador de iones pesados, del Laboratorio Lawrence de Berkeley, ambos en los Estados Unidos.

Leonardo Mascheroni, del Laboratorio de Los Alamos, se acerca a la postura de Rebut al pronunciarse sobre el proyecto internacional de reactor termonuclear. Mascheroni sostiene que deducir que pulsos de 10 millones de joule producirán la microfusión es una interpretación demasiado optimista de los datos de Centurion-Halite. Si el cebador no puede producir 10 veces esta energía, nunca desatará una ignición completa. La forma más económica de producir pulsos de 100 millones de joule, sigue afirmando, es con láseres de hidrógeno-flúor.

Queda aún otro problema: cómo conseguir una radiación uniforme del blanco para producir la ignición. Las técnicas se resumen en dos. La directa, que coloca múltiples haces de una forma simétrica alrededor del blanco; y la indirecta, que suspende el blanco en el centro de una cavidad, alineado con una lámina de oro o de otro metal pesado. Un haz de radiación entra a través de una abertura de la cavidad y vaporiza el metal, generando rayos X que convergen en el blanco. Por recurrirse a una técnica similar en las bombas termonucleares, los pormenores de la técnica se mantienen reservados, aunque sus partidarios afirman que las pruebas Centurion-Halite y otros ensayos obligan a suponer que se trata de la forma óptima de conseguir la ignición.

Disienten de ello los partidarios de la técnica directa. Aunque reconocen las dificultades existentes para comprimir blancos uniformemente, argumentan que los avances recientes pueden igualar los resultados conseguidos mediante la técnica indirecta. Señalan, además, que la técnica directa, al no requerir conversión de la fuente primaria de radiación en rayos X, dirige mejor la energía y, por consiguiente, reduce las necesidades de potencia del aparato en cuestión.

Una vez hecha la paz entre ambos bandos, habrá que hacer frente a otros problemas. Supongamos que Mascheroni esté equivocado y que un elemento cebador de 10 millones de joule baste para provocar la ignición; supongamos también que ésta produce 100 veces más energía que la radiación activadora. Pues bien, entonces cada encendido produciría mil millones de joule de energía, la misma energía producida cuando estallan 180 kilogramos de TNT. El laboratorio de microfusión generaría probablemente sólo uno de estos episodios cada día, pero al generador real se le exigiría realizarlo cinco o seis veces por segundo. ¿Qué recinto podría soportar estos violentos y frecuentes choques? Alguien ha propuesto que se rodee el núcleo del reactor con "cascadas" de litio licuado. Ahora bien, lo mismo que sucedía con el confinamiento magnético aquí importa la cuestión de escala. Un reactor que genere, por ejemplo, cinco mil millones de joule por segundo dejaría pequeño el mismo generador propuesto por Rebut.

Ciencia y empresa

Superordenadores

Se había convertido en hábito. Si alguien quería el ordenador más rápido, se encaminaba hacia Chippewa Falls, en el estado de Wisconsin, y firmaba el trato con Seymour Cray. Cray, una leyenda en su propia época, sigue en la brecha. Proyecta y construye superordenadores, máquinas que valen millones de dólares y resuelven complicados problemas numéricos antes que cualquier otro ingenio. En 1986, diez años después de que el primer superordenador saliera de su fábrica, Cray cifraba sus beneficios en un centenar y pico de millones de dólares y había instalado más de 120 máquinas.

Su empresa, Cray Research, vende el 80 por ciento de los superordenadores por doquier, salvo en el Japón. Los nipones son, precisamente, los que están promoviendo una creciente competencia con máquinas cuya velocidad de cálculo son comparables a las celérrimas de Cray. Además, los orientales producen sus propios componentes, son expertos en técnicas modernas de manufactura y tienen la suficiente resistencia financiera para soportar los altibajos de la demanda de superordenadores. Solamente IBM, que ha vuelto al campo de los superordenadores, viene con un empuje parecido.

Además, fuera de un puñado de compañías aeroespaciales, petrolíferas y automovilísticas, la industria se ha mostrado más cautelosa que entusiasta ante el futuro de los superordenadores. Los compradores se inclinan por un soporte lógico amplio, sistemas de uso sencillo y redes de comunicación de datos para conectar al superordenador múltiples investigadores. "Hemos de abandonar, afirma John A. Rollwagen, presidente de Cray Research, el objetivo de antaño, limitado a construir los equipos más rápidos, y lo hemos de hacer sobre la marcha y en plena jungla. Los equipos más rápidos son condición necesaria, pero no suficiente para el éxito."

La construcción de las máquinas más rápidas exige la técnica más depurada y el soporte lógico más refinado. Los elementos de memoria de alta prestación, los procesadores y los demás componentes deben colocarse apretadamente, a fin de reducir las distancias que las señales han de recorrer; los algoritmos que deciden cómo funciona la máquina y los conjuntos de programas para aplicaciones deben estar escritos de modo que redunden en la máxima eficacia.

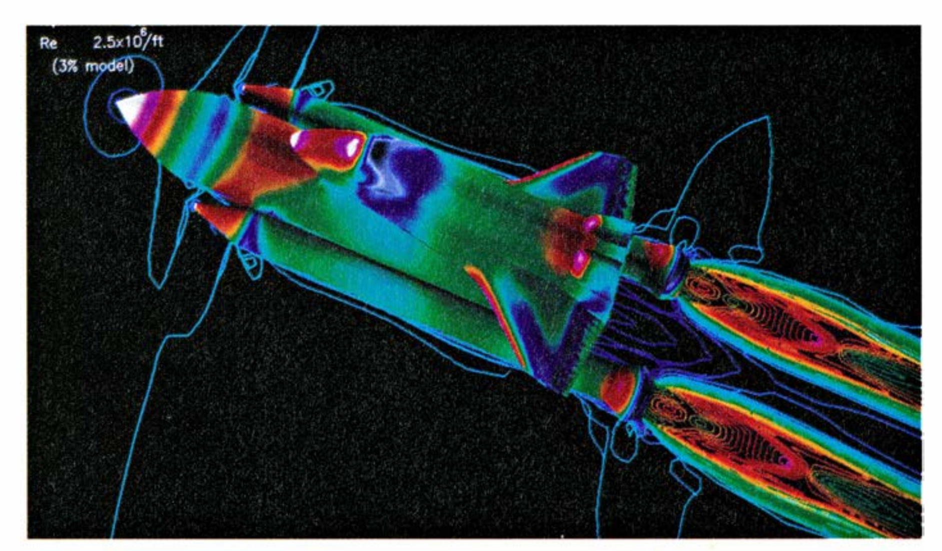
Fabricantes norteamericanos, encabezados por Cray, y competidores japoneses han seguido caminos diferentes hasta ahora. Los primeros aumentan la velocidad de sus máquinas enlazando múltiples procesadores que operan simultáneamente sobre diferentes partes de un problema. El CRAY-3, por ejemplo, que se entregará a los clientes en 1990, empleará hasta 16 procesadores, lo que le permitirá alcanzar velocidades de proceso de 16.000 millones de operaciones en coma flotante por segundo. Sin embargo, para aprovechar al máximo tales velocidades, los usuarios deben aprender a segmentar los problemas y así los procesadores múltiples puedan abordarlos. Por su parte, los actuales superordenadores japoneses se basan en un procesador "vectorial" combinado con la construcción y funcionamiento del equipo, que automáticamente busca operaciones idénticas en un programa y las calcula aproximadamente al mismo tiempo.

El soporte lógico utilizado por unos difiere del empleado por otros. Los norteamericanos emplean sistemas operativos basados en un conjunto de instrucciones y archivos ampliamente usado, llamado UNIX. Gran parte de los conjuntos de programas de aplicación se basa en el UNIX, lo que posibilita (aunque no de forma inmediata) su aplicación a máquinas distintas. En la trinchera antagónica, Fujitsu, NEC e Hitachi han construido superordenadores con sistemas operativos propios, compatibles con los modelos al uso. Esta política ayuda a sus clientes a pasar fácilmente del ordenador al superordenador, según explica Kenichi Miura, vicepresidente de Fujitsu America. Pero también ha traído como consecuencia que las máquinas japonesas no puedan trabajar con UNIX. Los empresarios japoneses comienzan, sin embargo, a ensayar los sistemas operativos basados en el UNIX.

Los superordenadores japoneses emplean un campo de números más limitado que los CRAY, aunque pueden expresar los números con mayor precisión. La capacidad de trabajar con números muy grandes se torna decisiva cuando los problemas adquieren complejidad creciente.

El carácter plural de las compañías japonesas del sector implica que ellas mismas fabrican los componentes necesarios para sus máquinas. Cray y otras empresas occidentales dependen de proveedores independientes —entre los que se encuentran sus competidores japoneses— para obtener los componentes necesarios. Steve S. Chen, antiguo ingeniero-jefe de Cray Research, que fundó su propia compañía de superordenadores en 1987, y que ha recibido apoyo de IBM desde entonces, se queja de esa situación que les deja a merced del competidor.

La carencia de técnicas avanzadas de fabricación constituye otro obstáculo para Cray y puede también serlo para otras empresas menores. Los módulos



PRESION DEL AIRE EN SUPERFICIE sobre el transbordador espacial, un minuto no cumplido después del lanzamiento; se ha calculado con el superordenador CRAY-2 del Centro de Investigación Ames de la Administración Nacional de Aeronáutica y del Espacio (NASA) en Moffet Field, California. Los cálculos, basados en un millón de puntos con datos, tardaron 20 horas en ejecutarse

de los procesadores del CRAY-3 medirán 100 centímetros cuadrados y estarán conectados a otros módulos por 12.000 conexiones; sólo una línea de montaje robotizada puede fabricar el sistema con precisión. No se olvide que esas dificultades en establecer la línea de montaje contribuyeron a retrasar el lanzamiento del CRAY-3 en más de seis meses.

Además de incomodar a los clientes, los retrasos de producción se notan sobre todo en los mercados de valores. A fines de octubre, las acciones de Cray cayeron aproximadamente un 17 por ciento, en cuanto los analistas se enteraron no sólo de una disminución en los beneficios trimestrales, sino también de retrasos de producción y dificultades en el desarrollo de los microtransistores lógicos del CRAY-3, lo que significaba que la compañía tenía que gastar 10 millones adicionales de dólares en su desarrollo.

Mientras Cray y otros fabricantes han de habérselas con esos problemas, muchos usuarios de superordenadores se preguntan si realmente importa qué empresa —o país— fabrica los ordenadores más rápidos. Las ventas no alterarán el déficit comercial de Estados Unidos. En 1988, las ventas mundiales de superordenadores se acercaron a 1000 millones de dólares: menos de una cuarta parte de las ventas anuales de entradas de cine en los Estados Unidos. "Los beneficios de los superordenadores no provienen de la creación de máquinas, sino de su uso", comentó el senador por Tennessee Albert Gore Jr., en un simposio celebrado, el otoño pasado, por la Academia Nacional de Ciencias.

La propia velocidad de proceso de los superordenadores, más de 1000 veces la de un potente ordenador personal, posibilita que los investigadores ataquen complicados problemas, inabordables sin su auxilio. Citemos, a modo de ejemplo, el modelado de la corriente de aire en torno al ala de un avión, el descubrimiento de pozos de petróleo, basándose en montones de datos sísmicos, o el cartografiado del genoma humano. Además, las máquinas pueden convertir una avalancha de datos en una imagen tridimensional en falso color, capaz de espolear a los investigadores hacia nuevas cuestiones. Los superordenadores son "máquinas de la imaginación", dice Sidney Karin, director del centro de supercálculo de la Universidad de California en San Diego. Con tal que los investigadores tengan acceso a las máquinas más rápidas, añade Karin, no importa dónde se construyan.

El equipo físico de un superordenador de primera categoría cuesta 20 millones de dólares, lo que significa que la decisión de adquirir uno no puede tomarse a la ligera. "Nuestra mayor dificultad estriba, sin embargo, en hallar la forma de promover el uso de los superordenadores", dice Clifford R. Perry, director de investigación de Eastman Kodak. Confusos ante los superordenadores, los investigadores prefieren seguir trabajando con ordenadores que ya entienden, apostilla. Asociarse con centros universitarios de superordenadores puede ser una buena manera de que quienes trabajan en empresas aprendan el manejo de estas máquinas, insinúa el vicepresidente de Amoco. A pesar de todo, el éxito se logrará el día en que los programas funcionen de la misma manera en un ordenador de mesa que en un superordenador; entonces, usuarios de los cuatro puntos cardinales, conectados por una red, podrán extraer toda la potencia de un superordenador anónimo.

Aguzando el filo

La batalla de las latas se está cociendo en el refrigerador. En los años sesenta, las latas de cerveza y bebidas refrescantes se hacían de hojalata. Veinte años más tarde, el aluminio ha ocupado su lugar. Pero en el último bienio las mejoras introducidas en los procesos de manufactura de hierro y acero, junto con la subida de los precios del aluminio, han animado a las siderúrgicas a recoger el guante que el aluminio les había echado por la primacía en el mercado de bebidas enlatadas.

La vieja lata de bebidas era un trasto pesado, con tendencia a rezumar, que constaba de tres piezas soldadas; además, sazonaba la bebida con un sabor metálico. Los fabricantes de aluminio desarrollaron luego un proceso que permitía producir un recipiente a partir de una lámina plana; sólo faltaba la tapa. Había una objeción seria contra las latas de aluminio: no se degradaban por oxidación cuando se echaban en vertederos. Pero la industria salió airosa con una gran campaña de reciclado, al mismo tiempo que atraía a los fabricantes de latas hacia el aluminio.

Sin embargo, en los últimos cinco o seis años, las siderúrgicas que sobrevivieron a la crisis empezaron a remozar sus fábricas. En el proceso de antaño, las laminadoras comenzaban con barras de acero, las convertían en láminas de dos décimas de milímetro de espesor y entonces las estañaban. Ahora, los mismos fabricantes usan un proceso de fundición continua que produce largas tiras de hierro incandescente que después se laminan hasta espesores de 0,07 milímetros. El espesor de la capa de estaño se ha reducido a la mitad. Esta hojalata más fina se puede convertir en recipientes cilíndricos sin costura, lo mismo que el aluminio. Tales latas tienen, no obstante, la tapa de aluminio: no se puede hacer todavía una tapa de hojalata que se abra tirando simplemente de la anilla. Pero los siderúrgicos saben que el aumento del precio del aluminio labora en su favor.

No son todavía muchos los que se han pasado a la hojalata. Hay que reequipar las líneas de producción; adaptar una planta de aluminio a la producción de hojalata puede costar hasta 350 millones de pesetas. Un obstáculo mayor con el que se enfrentan los siderúrgicos es el del reciclado, un proceso muy costoso en la hojalata. Por ser de alta calidad las finas láminas de hierro que se usan para las latas, los fabricantes deben vigilar cuidadosamente el nivel de otros metales, en particular del estaño, en el hierro reciclado, para vigilar que no se torne quebradizo. En muchos casos resulta más caro reciclar un recipiente de hojalata que hacer uno nuevo.

Aun cuando las pequeñas siderúrgi-

cas, el sector de más rápido crecimiento de la industria del acero, emplean chatarra, han carecido hasta ahora de los grandes trenes de laminación necesarios para crear chapas. Aunque las cosas están cambiando. Dentro de unos meses, una empresa norteamericana abrirá una planta que usará la técnica de fundición continua ideada en Alemania occidental para fabricar láminas de hierro; producirá al principio chapas de 2,5 a 6,5 milímetros de espesor, pero espera con el tiempo refinar el proceso y producir chapa de hierro de alta calidad.

Un camino a seguir es el de la unión. En Estados Unidos, los seis principales productores de hojalata han acordado reunir 30 millones de dólares en los próximos cinco años para el reciclado conjunto de los recipientes de hojalata. El organismo creado desarrollará redes locales de reciclaje, tomando ejemplo de un programa similar desarrollado con éxito en Ontario (Canadá).

Apareció un ratoncillo

Ha nacido un nuevo ratón en la granja de roedores de Río Charles. La verdad es que pronto vendrán por centenares los nuevos ratoncillos genéticamente idénticos en la granja, que cría roedores para sus clientes. Se trata de los "oncorratones", fabricados, en cierto sentido, por DuPont, la vieja compañía de productos químicos.

Los oncorratones son las criaturas que saltaron de las revistas científicas a la primera página de los periódicos en el mes de abril, cuando la Oficina de Patentes y Marcas Registradas de los Estados Unidos concedió a la Universidad de Harvard la primera patente referente a un animal transgénico. Se espera que tales ratones, genéticamente predispuestos a desarrollar cáncer, posibilitarán que los investigadores sometan a ensayo nuevas terapias y —por ser mejores modelos de la enfermedad— reduzcan el número de cobayas.

La nueva raza fue desarrollada por Philip Leder y Timothy A. Stewart, de Harvard. Implantaron un oncogén humano en un embrión de ratón al poco de su concepción. Leder realizó el trabajo, subvencionado en buena parte; en contrapartida, la empresa química se reservaba los derechos comerciales sobre los resultados de la investigación.

Lo que menos se esperaba DuPont era un ratón patentado. "El día que me enteré de la patente", recuerda un alto representante de la empresa, "me pregunté: ¿qué negócio vamos a sacar de esto?".

Transcurrieron seis meses e hizo falta una grave caída de las ventas para convencer a la dirección de DuPont de que los ratones encajarían cómodamente entre las líneas de productos de la compañía. Aceptado el reto, la compañía se llevó cuatro ratones del laboratorio de Leder y comenzó a criar.

"Leder inventa y nosotros retocamos", dice Forman. DuPont empezó con oncorratones heterozigóticos (ratones que llevan una copia del oncogén cada uno) y cruzaron los roedores para producir una población de ratones homozigóticos (que llevan dos copias del oncogén) confirmados. Río Charles se encargó de ellos a partir de enero. Para producir ratones para la venta, hay que cruzar el linaje endógamo de machos homozigóticos con un linaje igualmente endógamo de hembras normales. La endogamia intensiva da por resultado crías idénticas y heterozigóticas.

En Río Charles, los ratones residen en habitaciones esterilizadas, protegidas por ordenador de posibles contaminantes. La compañía conserva incluso algunos linajes de ratones congelando embriones en nitrógeno líquido para el caso de que una línea experimentase una mutación espontánea. Cuando los ratones salen de la granja, viajan en cajones especiales transportados por camiones climatizados.

DuPont ofrecerá los ratones inicialmente a grupos no lucrativos y universidades. ¿El precio? Unos 50 dólares por ratón, con descuentos para grandes pedidos. Río Charles dice que el precio es competitivo frente a otros ratones de especialidad; las variedades comunes se venden a dólar por animal.

No hay nada, concede Forman, que impida a los investigadores criar oncoratones en sus propios laboratorios. Aunque DuPont tomó en consideración la posibilidad de vender solamente animales de un sexo o esterilizar los ratones como modo de proteger sus derechos de patente, decidió confiar en el código de honor y en el hecho de que los investigadores frecuentemente necesitan estar seguros de que sus ratones son genéticamente uniformes.

Como la patente cubre técnicamente todos los mamíferos transgénicos no humanos que lleven oncogenes, y hay unos 40 oncogenes conocidos, Forman predice que el actual ratón será sólo el primero. La compañía escribirá a los investigadores que hayan producido sus propios oncorratones pidiéndoles que paguen unos derechos nominales de usuario de la licencia y solicitando si desearían colaborar con el proyecto de los ratones de DuPont.

Mezcla de fluidos

Los fluidos viscosos que fluyen según esquemas sencillos y periódicos en dos dimensiones pueden generar el caos que conduce a una mezcla eficiente. Los experimentos y los modelos de ordenador sacan a la luz el mecanismo subyacente

Julio M. Ottino

¿Qué tienen en común la erupción del Krakatoa, el amasijo de la harina y el brillo de las estrellas? Que en todos se da algún tipo de mezcla. La mezcla violenta de los magmas podría haber desencadenado la erupción del Krakatoa; al estirar y doblar —el procedimiento clásico de mezcla— es lo que requiere el amasijo, y, por último, la mezcla en el interior de una estrella determina su composición química y, en consecuencia, el brillo de su superficie. El universo está lleno de ejemplos de mezclas, que abarcan todas las escalas de tiempo y longitud. Los gases de un escape se mezclan con el aire exterior en cuestión de segundos, mientras que el proceso de mezcla en el manto de la Tierra necesita cientos de millones de años, si no más.

El fenómeno de la mezcla desempeña también un papel crítico en la tecnología moderna. La ingeniería química se basa en la mezcla para asegurar que las sustancias reaccionen adecuadamente, para producir combinaciones de polímeros que presenten propiedades singulares y para dispersar agentes que reduzcan el flujo en las tuberías. Sin embargo, a pesar de su ubicuidad en la naturaleza y en la industria, queda mucho por conocer del proceso de mezcla. Los mismos investigadores ni siquiera cuentan con una terminología común: mezclar es sinónimo de "remover" para los oceanógrafos y geofísicos; mezcla significa "composición" en el argot de los ingenieros de polímeros y, "agitación", para los ingenieros de procesos.

Con independencia de cómo se denomine el proceso, lo incuestionable es su extrema complejidad y el hallarse en sistemas muy dispares. En la construcción de una teoría de mezcla de fluidos, hay que tener en cuenta los fluidos y los flujos; aquéllos pueden ser miscibles o parcialmente miscibles y pueden reaccionar o ser inertes; los flujos, por su lado, son lentos y ordenados o muy rápidos y turbulentos. Por tanto, no es sorprendente que ninguna teoría única pueda explicar todos los aspectos del proceso de mezcla en los fluidos y que los cálculos directos sean normalmente incapaces de tener en cuenta todos los detalles importantes.

A pesar de todo, tanto los experimentos físicos como las simulaciones mediante ordenador pueden proporcionar indicaciones acerca del proceso de mezcla. Durante los últimos años, mis colaboradores y yo hemos seguido ambas aproximaciones, en un esfuerzo por dilatar nuestros conocimientos sobre varios aspectos del proceso; en particular, los concernientes a mezclas que involucren flujos lentos y fluidos viscosos, como los aceites.

Remover dos pinturas oleosas es un buen ejemplo de mezcla de fluidos viscosos. Después de agitar unos segundos, se produce un motivo complejo de líneas que se estiran y se doblan. (Los encuadernadores aprovechan ese recurso para crear los "jaspeados" que a veces adornan las cubiertas o contracubiertas de los libros.) Ahora bien, a no ser que hayamos agitado a conciencia, quedarán, probablemente, algunas "islas" de pintura sin mezclar entre las líneas revueltas. Aunque la mezcla de fluidos viscosos puede producir estructuras fantásticamente complicadas, puede producir también motivos o esquemas que presentan cierta regularidad y coherencia.

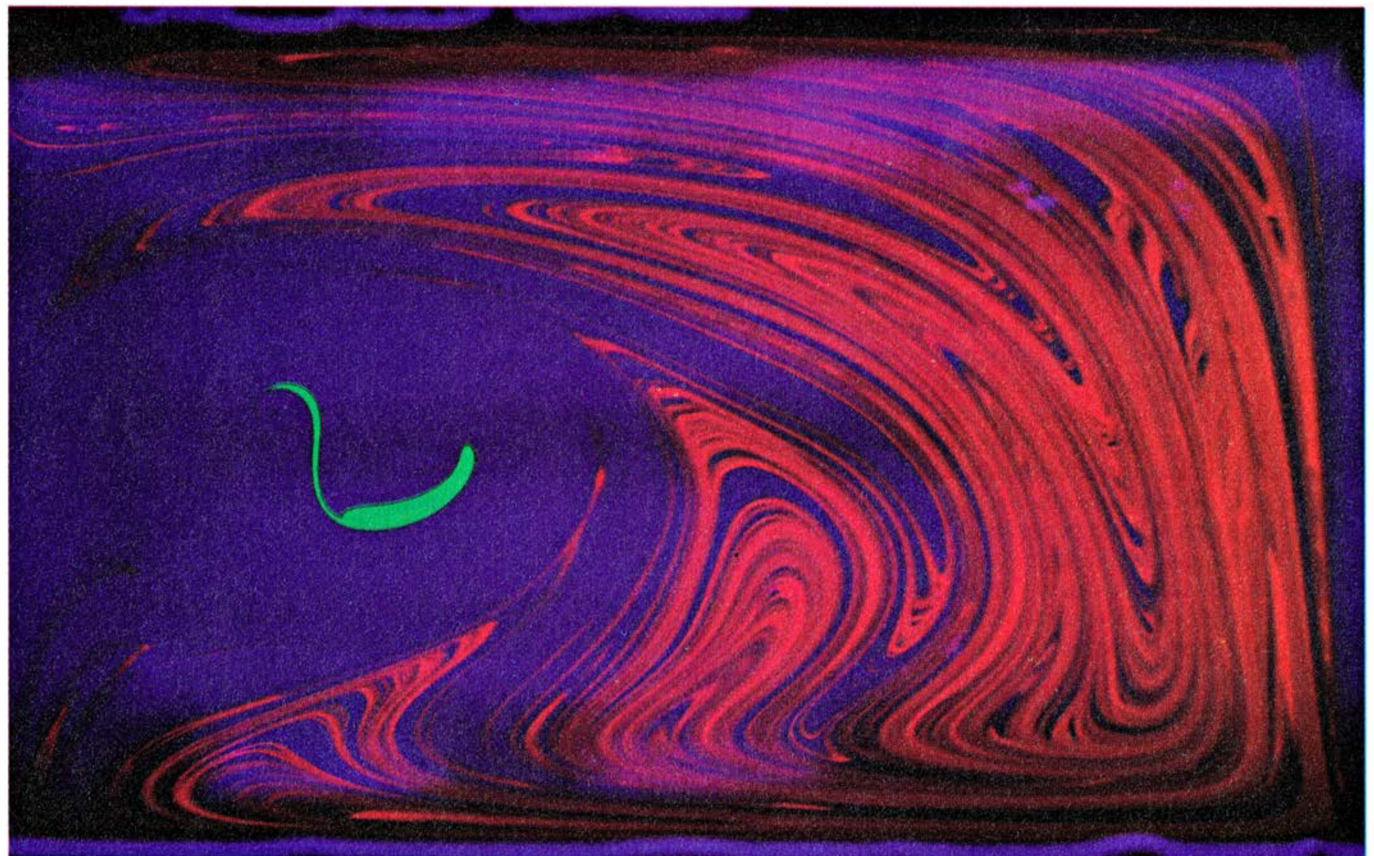
En la Universidad de Massachusetts en Amherst, he intentado, con mis alumnos, caracterizar los flujos que producen tales esquemas. Realizamos con ese propósito experimentos y simulaciones mediante ordenador que remedaran la mezcla de dos pinturas. En algunos experimentos, inyectamos gotas de glicerina coloreada en un volumen de glicerina incolora contenida en un recipiente profundo. Cuando se mueven de forma periódica las paredes del recipiente, los esfuerzos que éstas desarrollan sobre el fluido viscoso de la vasija pueden estirar y doblar la gota coloreada de una manera bastante complicada; el recipiente entero presenta pronto un complejo esquema de pliegues dentro de otros pliegues. Sin embargo, una gota similar del mismo recipiente puede, con idéntica facilidad, no experimentar apenas deformación; la gota puede moverse y girar, pero vuelve regularmente a su posición inicial. ¿Cómo surgen motivos o creaciones tan dispares?

Nociones de mecánica de fluidos

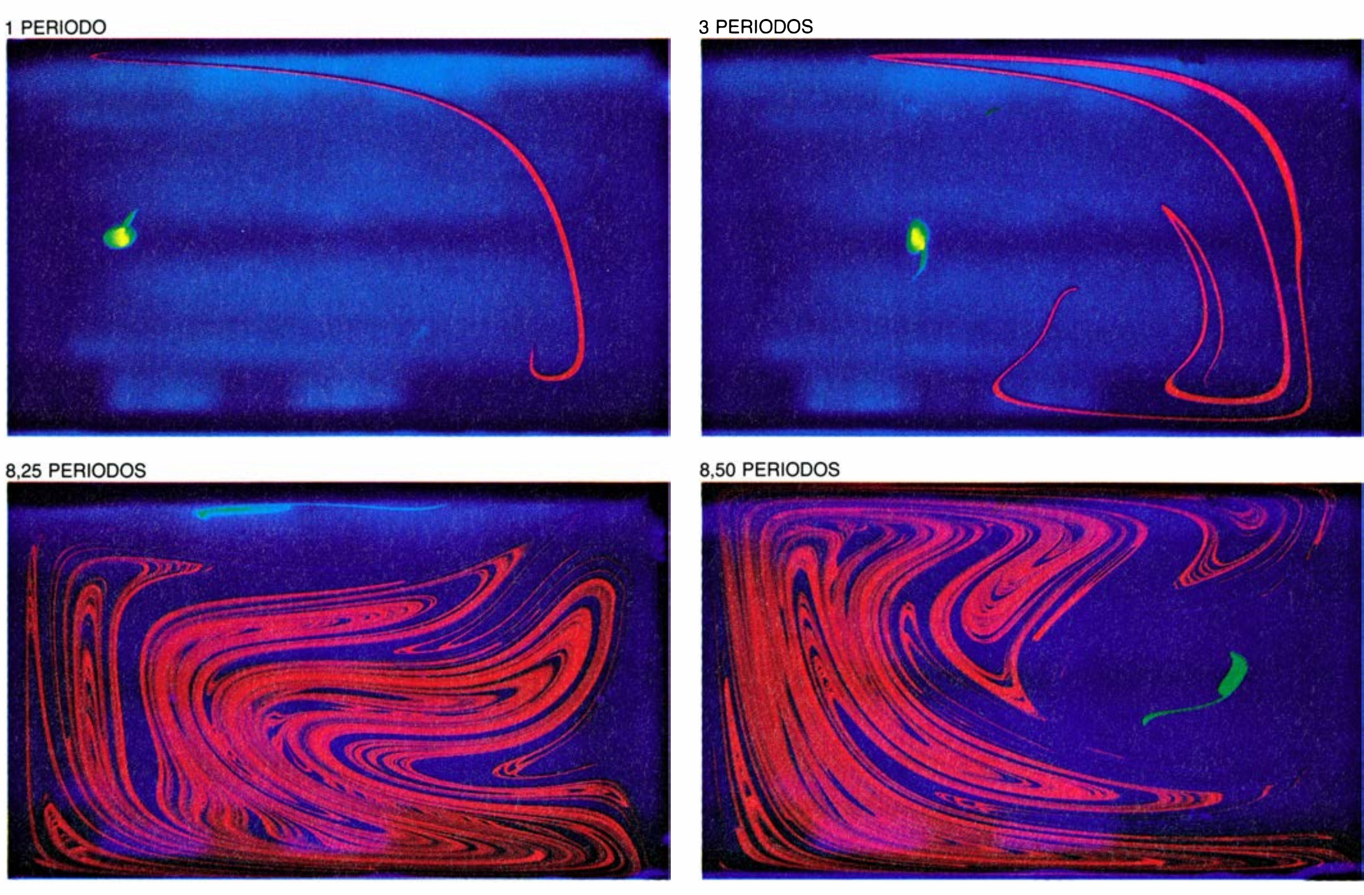
La clave para entender los aspectos básicos de tales mezclas reside en la idea de "movimiento", concepto que se remonta al matemático suizo del siglo XVIII Leonhard Euler. El movimiento de un fluido es una expresión matemática que establece dónde se hallará cada partícula del fluido en cualquier instante futuro. Si se pudiera conocer el movimiento para un flujo determinado, se podría averiguar, en principio, cuanto se precisa sobre la mezcla que produce. Se podrían calcular, por ejemplo, las fuerzas y la energía total necesarias para lograr cierto grado de mezcla en el sistema.

Durante el siglo pasado, la descripción del flujo en función del movimien-

JULIO M. OTTINO es profesor de ingeniería química y de polímeros en la Universidad de Massachusetts en Amherst. Estudió en la Universidad Nacional de La Plata y en la de Minnesota, donde obtuvo su doctorado en 1979. Su experiencia en las mezclas de materiales coloreados no se limita al laboratorio, sino que se prolonga en su labor pictórica, que se exhibe ahora en una exposición en Argentina.



1. FLUJOS CAOTICOS Y NO CAOTICOS, representados en contraposición, según un experimento realizado por Kenny Leong y el autor en su laboratorio de la Universidad de Massachusetts en Amherst. Llenamos de glicerina un recipiente rectangular e inyectamos dos gotas de sustancias trazadoras fluorescentes, respectivamente, justo debajo de la superficie (*arriba*). Cada lado de la cavidad puede deslizarse en una dirección paralela a ella con independencia de los otros lados. En este caso particular, se hicieron mover periódicamente, aunque de manera discontinua, los lados de arriba y abajo. El lado de arriba se movía de izquierda a derecha y luego se detenía, en el momento mismo que el lado de abajo empezaba a moverse a la misma velocidad y durante el mismo rato, si bien de derecha a izquierda; este par de movimientos constituye un período. Transcurridos 10 períodos de ésos (*abajo*), la gota roja se había estirado y plegado varias veces: se había colocado en una zona de mezcla caótica. La gota verde sólo se había estirado un poco: era una ''isla'' de mezcla no caótica.



2. ESTIRADO Y PLEGADO característicos de las mezclas caóticas; ambos procesos se indican mediante la gota roja en este conjunto de fotografías del experimento anterior. Al término de los tres primeros períodos se distingue ya con nitidez el proceso básico de estirar y plegar. La isla verde que indica una región de mezcla principalmente no caótica y los pliegues que indican una región de mezcla caótica se mueven por el recipiente, pero vuelven a sus posiciones

to de un fluido ha quedado muy relegada en favor de una descripción basada en el campo de velocidades del fluido: una expresión que especifica la velocidad del fluido en cada punto de la zona de movimiento y en cada instante de tiempo. Si se conoce el movimiento, es fácil calcular el campo de velocidades, mientras que el conocimiento del campo de velocidades no basta para calcular explícitamente el movimiento. El movimiento es, por tanto y en cierto sentido, una descripción más fundamental del flujo; en virtud de ello, mis colaboradores y yo preferimos trabajar con lo que muchos reputarán tal vez un concepto anticuado.

Subyacente al concepto de movimiento, encontramos la transformación de punto. Se trata de una operación matemática que permite identificar una partícula de un fluido y determinar su posición en cualquier instante futuro. Cada partícula del fluido se "representa" en una nueva posición mediante la aplicación de la transformación. Las partículas inicialmente identificadas como separadas no pueden ocupar la misma posición en el mismo instante, ni una partícula puede desdoblarse en dos. Aunque, en teoría, existe una transformación de punto para todos los flujos de mezcla, sólo en los casos más sencillos puede obtenerse exactamente. Por esta razón, mucho de lo que se conoce acerca del proceso de mezcla se limita a flujos de relativa sencillez, verbigracia, flujos lineales en los que las líneas de un trazador no se doblan. Estos tipos de flujos quizá no abarquen los procesos que llevan a mezclas eficientes y que son inherentemente no lineales. Para tener, por lo menos, una idea de qué se encierra en esos procesos, hay que considerar flujos estacionarios en dos dimensiones.

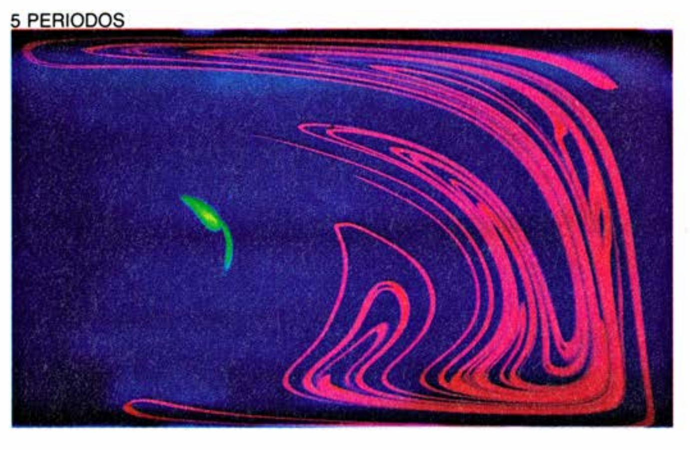
Flujos en dos dimensiones

Todos los flujos bidimensionales constan de las mismas piezas fundamentales: puntos hiperbólicos (o de silla) y puntos elípticos [*véase la figura 4*]. Un fluido se mueve hacia un punto hiperbólico en una dirección y se aleja de él en otra, mientras que un fluido circula en torno a un punto elíptico. (Importa mencionar un tercer tipo de punto, llamado parabólico, en el que el movimiento del fluido es de cizalla o tangencial. Es el tipo que encontramos en el fluido que pasa a lo largo de una pared sólida. Los puntos parabólicos pueden olvidarse al describir la naturaleza de las mezclas en flujos bidimensionales.) Según cabría esperar, la mezcla en un flujo bidimensional estacionario es bastante ineficaz en comparación con la mezcla en flujos tridimensionales, en especial en aquellos que cambian continuamente con el tiempo. La verdad es que hay sólo dos posibilidades principales en un flujo bidimensional estacionario y acotado: o las partículas del fluido siguen repetidamente las mismas trayectorias, conocidas como líneas de corriente, o bien no se mueven en absoluto.

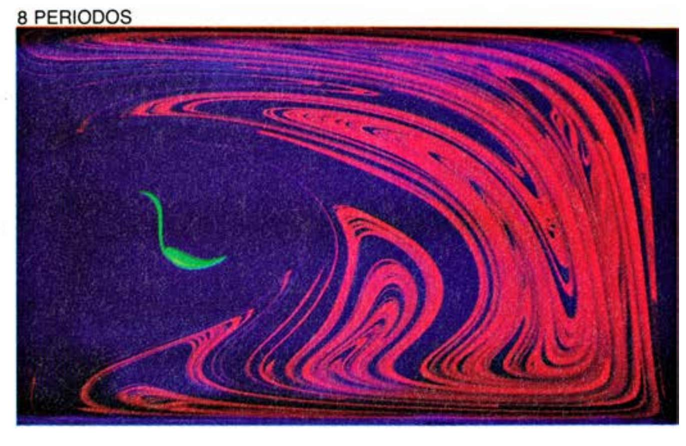
Las líneas de corriente en los fluidos estacionarios son fijas y las trayectorias de las partículas del fluido nunca se cruzan; ello implica que las partículas del fluido no disfrutan de la oportunidad de entrar en contacto entre ellas, es decir, de mezclarse. ¿Existe alguna manera de salirse del confinamiento de las líneas de corriente, de manera que las partículas del fluido se libren de tener que seguir repetidamente la misma línea de corriente del flujo? La hay si el esquema del flujo puede cambiar con el tiempo, de suerte tal que una línea de corriente de un esquema se cruce con una línea de corriente de otro esquema posterior.

La forma más sencilla de realizarlo (y la más fácil de analizar) es la de obligar al flujo a variar, de manera periódica, con el tiempo. Sin embargo, para

5 PERIODOS

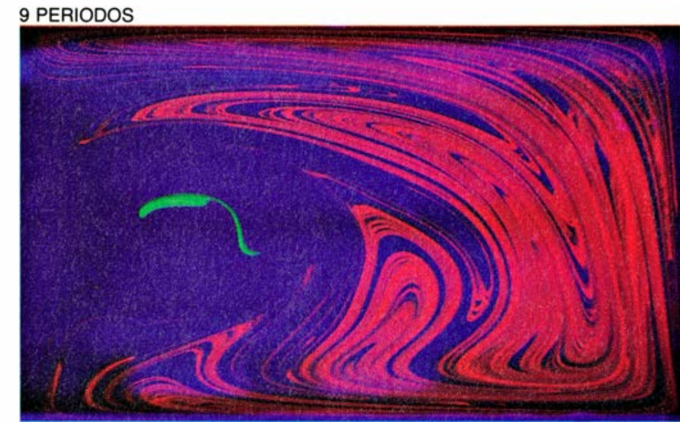


8 PERIODOS



8,75 PERIODOS

9 PERIODOS



(aunque algo deformadas) después de cada período. El recorrido que ha dibujado la gota verde señala que ésta experimenta una rotación completa cada dos períodos. Si el experimento se proyectara atrás en el tiempo, la gota verde volvería más o menos a su forma y posición iniciales, ya que el error en la reproducción de sus movimientos, al ir hacia atrás, crece linealmente. El error en la reproducción de los movimientos inversos de la roja crece de manera exponencial.

que tal flujo conduzca a una mezcla eficiente, debe ser capaz de estirar y plegar una región del fluido y devolverla —estirada y plegada— a su ubicación inicial. La operación de estirar y plegar corresponde a una aplicación en herradura, según la describe Stephen Smale, de la Universidad de California en Berkeley.

Resulta contrario a la intuición que, para mezclar bien un material, debamos devolver parte del mismo a su posición inicial. Ahora bien, si la mezcla se produce en un sistema acotado, no queda otra alternativa. Si lanzamos una y otra vez dardos contra una diana, alguno acabará por caer tan cerca como se quiera de cualquier otro, por la sencilla razón de que la superficie de la diana es finita. Del mismo modo, estirando y plegando reiteradamente en el interior de un recipiente cerrado, habrá momentos en que las partículas de fluido tornen a hallarse muy cerca de sus posiciones iniciales.

Si una partícula de fluido de un flujo periódico vuelve a su posición inicial exacta después de cierto tiempo, tal partícula define un punto periódico. Dependiendo del número de períodos necesarios para que la partícula vuelva a su posición inicial, se dice que es un punto periódico de período uno, período dos, etcétera. Un punto periódico puede clasificarse también como punto hiperbólico o punto elíptico, según sea la dirección del flujo en su entorno inmediato.

Conforme un punto periódico elíptico va dibujando su trayectoria cíclica, la materia que rodea al punto no sólo circula a su alrededor (tal como lo haría en torno a un punto elíptico fijo), sino que también se mueve con él. Mas, a pesar de la rotación y la traslación del material, no se desprende fácilmente materia hacia el resto del flujo. Tales regiones del material se ven como "islas" de fluido; en el interior de éstas, la mezcla suele desarrollarse con lentitud. Como la materia no puede entrar ni salir de la vecindad de un punto elíptico periódico, tales puntos entorpecen el rendimiento de la mezcla.

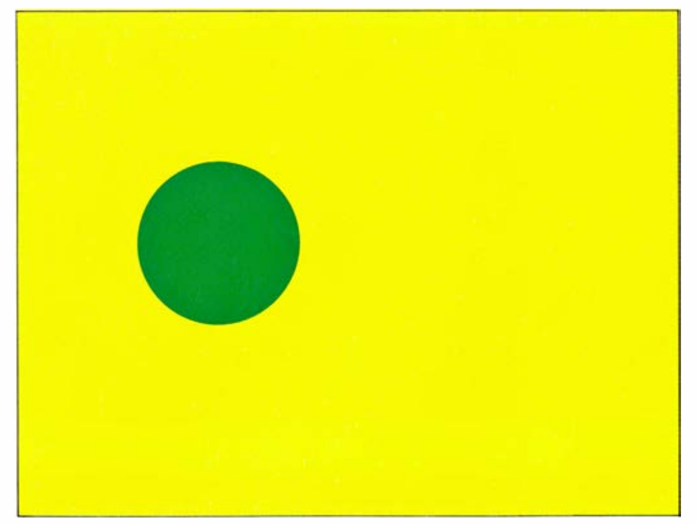
De manera similar, cuando un punto hiperbólico periódico describe su trayectoria cíclica, la materia que lo rodea y se mueve con el punto sufre una contracción en una dirección y se estira en la otra. Al hacerlo, el punto desprende filamentos estirados en una dirección y atrae material en la otra. (Si se supone que los fluidos son incompresibles, el estiramiento y la contracción deben equilibrarse entre sí.)

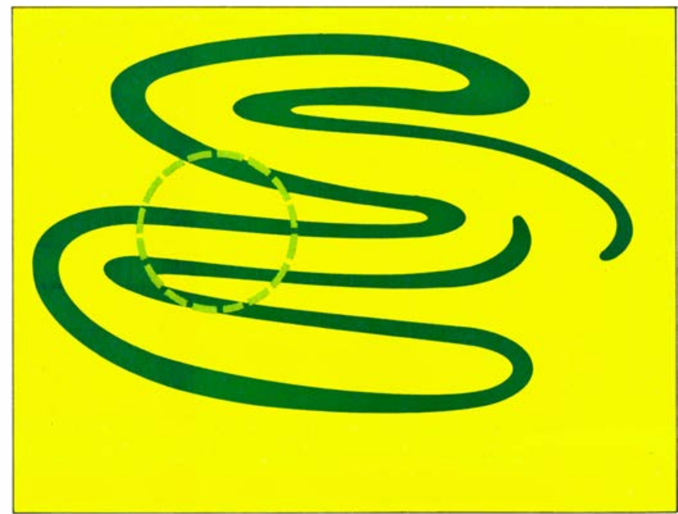
¿Dónde va la materia desprendida de un punto hiperbólico periódico? ¿De dónde viene la materia que se acerca al punto? Cabe la posibilidad de que un flujo entrante se una suavemente con un flujo saliente: que la materia que se desprende de un punto hiperbólico sea atraída por el mismo punto hiperbólico o por otro. Esto es lo que, de hecho, sucede en un flujo estacionario (aunque en dicho caso los puntos hiperbólicos son fijos y no periódicos); como resultado, el flujo no estira ni pliega la materia con la eficacia debida.

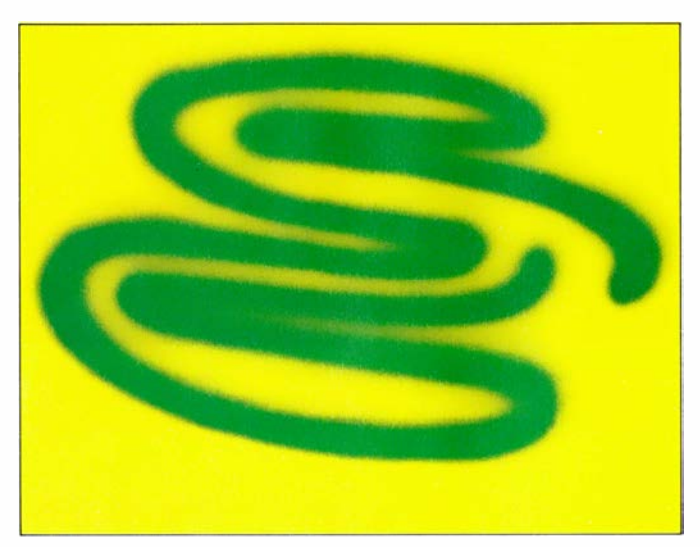
Huella del caos

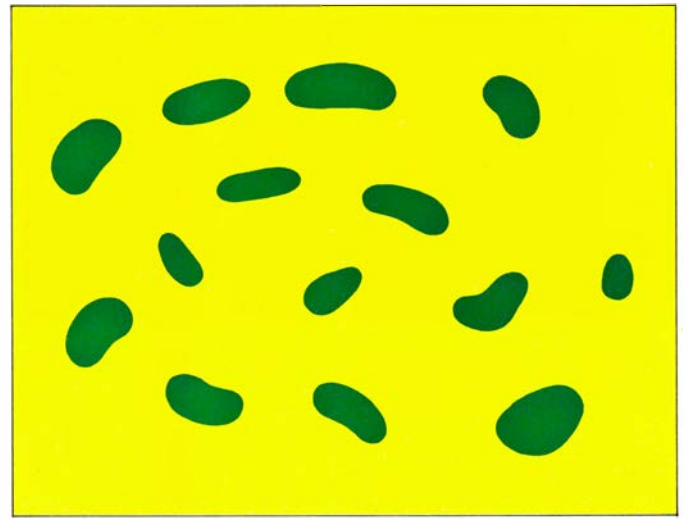
Sin embargo, los flujos bidimensionales que dependen del tiempo pueden crear mecanismos eficientes de estirar y plegar, gracias a que, en tales flujos, es posible que una región saliente y asociada a un punto periódico hiperbólico se cruce con la región de flujo entrante del mismo punto hiperbólico, o de otro. Se denomina punto homoclínico transversal aquel donde se cortan el flujo entrante y el flujo saliente de un mismo punto hiperbólico. Si la intersección se debe a los flujos de dos puntos hiperbólicos distintos, hablaremos de punto heteroclínico transversal.

Las intersecciones homoclínicas y heteroclínicas son las huellas del caos.









3. EN LA MEZCLA de fluidos, mediante procesos naturales o industriales, intervienen la difusión molecular y roturas, así como mecanismos de estiramiento y pliegue. Sólo en un caso ideal, una gota de tinte en un fluido puede (*arriba, a la izquierda*) estirarse y plegarse indefinidamente sin difundirse o romperse (*arriba, a la derecha*). Rasgo interesante de ese caso hipotético es el de que, para obtener una mezcla eficiente, parte de la gota debe volver a su posición inicial. La difusión molecular (que imposibilita siempre la mezcla perfecta) difumina corrientemente los límites entre fluidos miscibles (*abajo, a la izquierda*). En el caso de fluidos inmiscibles, la gota, al estirarse, puede dividirse en nuevas gotitas que, a continuación, pueden juntarse y formar varias gotas menores.

Desde un punto de vista matemático, un sistema que es capaz de producir mapas de herradura o intersecciones homoclínicas o heteroclínicas transversales puede clasificarse como caótico. De hecho, un mapa de herradura implica la existencia de puntos homoclínicos transversales; de la misma manera, la identificación de tan sólo un punto así basta para exigir la existencia de un mapa de herradura.

A Henri Poincaré, matemático francés de la segunda mitad del siglo XIX y principios del XX, debemos el descubrimiento de la observación según la cual un cruce de flujos entrantes y salientes conduce invariablemente a puntos homoclínicos transversales y de que tal cruce puede ocurrir incluso en lo que parecen ser sistemas físicos de "buen comportamiento", descritos por las leyes del movimiento de Newton. Pero el análisis del comportamiento, abrumadoramente complejo, que se deriva de tales cruces (lo que hoy llamamos caos), trascendía las fuerzas de Poincaré, quien decidió dejar así el asunto.

Mientras el proceso de mezcla pueda representarse mediante una transformación de punto determinista, la mezcla debe ser cinemáticamente reversible. En otras palabras, ha de ser posible "desmezclar" los fluidos (por lo menos, si se desprecia la difusión molecular). Ahora bien, la experiencia diaria sugiere que la mezcla constituye un proceso irreversible. A pesar de que el sistema es, en teoría, determinista, los movimientos que llevan a estirar y plegar repetidamente no pueden deshacerse.

Una situación bastante similar existe en otros sistemas físicos, como los analizados por Poincaré, que consisten en muchas partículas cuyos respectivos movimientos se describen mediante ecuaciones deterministas. (Estos sistemas suelen calificarse de hamiltonianos.) Uno de los físicos norteamericanos más notables del siglo XIX, J. Willard Gibbs, reconoció que hasta los sistemas hamiltonianos podían tener una irreversibilidad e impredictibilidad connaturales; que él propusiera un experimento teórico relacionado con las mezclas, a fin de explicarlo, nos da la medida de su intuición. Nadie pareció caer en la cuenta de su observación hasta que el oceanógrafo sueco Pierre Wellander destacó su interés en un perspicaz artículo de 1955.

Captura del caos en flujos

La importancia que reviste la acción de estirar y plegar en el proceso de mezcla se conoce, en ingeniería química, desde la década de 1950, a raíz del trabajo pionero de Robert S. Spencer y Ralph M. Wiley, de la empresa Dow Chemical, y de William D. Mohr y sus colaboradores de la firma du Pont de Nemours. Las consecuencias de este hecho —la existencia de mapas de he-

rradura y puntos homoclínicos y heteroclínicos— permanecieron ocultas hasta hace poco.

Se atribuye al matemático ruso Vladimir I. Arnold la primera conexión directa entre caos y flujos de fluidos. Según Michel Hénon, astrónomo francés del Observatorio de Niza, Arnold sugirió en 1965 la posibilidad de que los sistemas de mecánica de fluidos presentaran trayectorias caóticas. Hénon siguió la conjetura de Arnold y, en un trabajo de tres páginas que sólo contenía una figura, demostró que un flujo de fluido estacionario tridimensional sin viscosidad puede crear líneas de corriente caóticas.

En 1984, Hassan Aref, entonces en la Universidad de Brown, observó que las ecuaciones que describían las trayectorias de las partículas de un fluido en un flujo bidimensional eran formalmente idénticas a las que describían un sistema hamiltoniano. Progresó en su observación demostrando, mediante una simulación de ordenador, que un sistema hamiltoniano sujeto a fuerzas periódicas podía producir una mezcla eficiente.

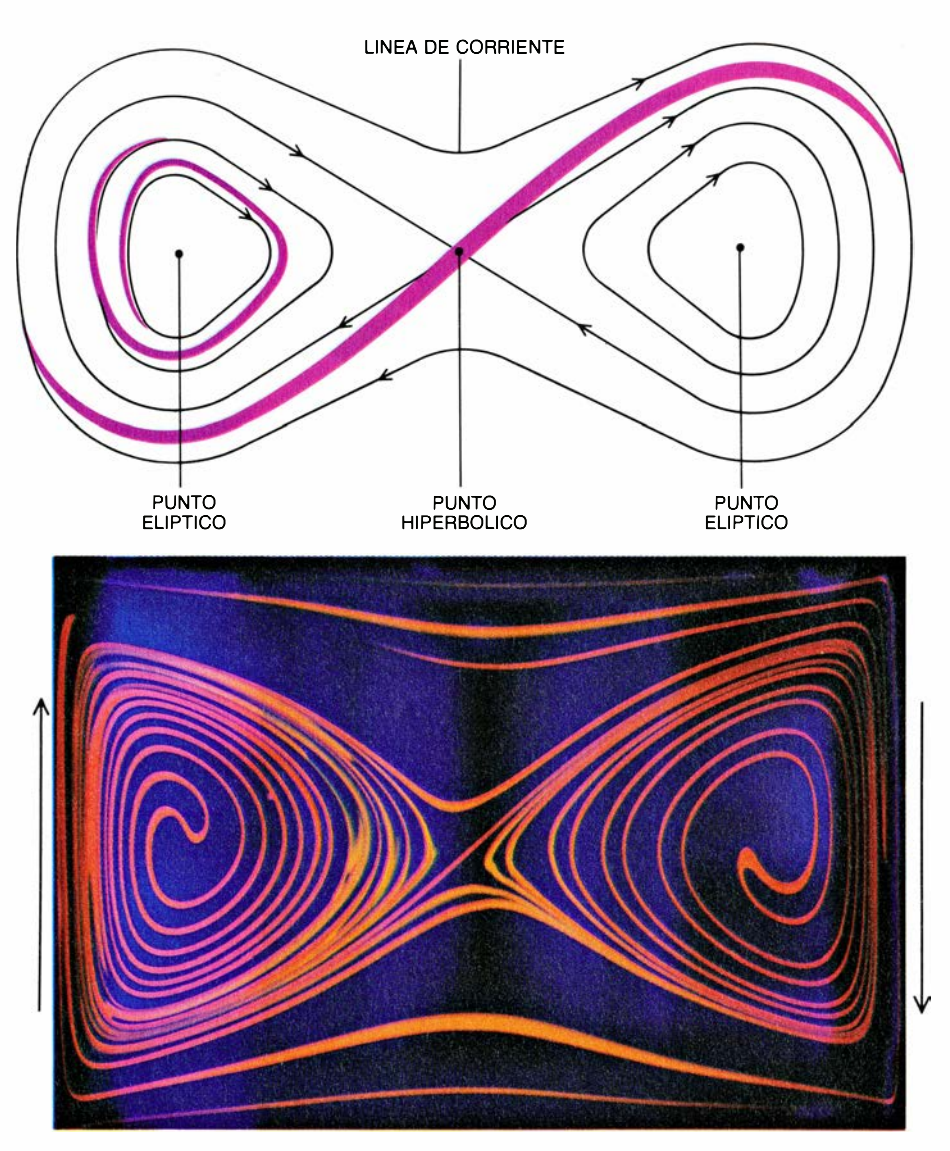
En tres dimensiones, la analogía entre el proceso de mezcla y los sistemas hamiltonianos no funciona; en dos, la analogía es exacta: la mezcla de fluidos puede considerarse representación visual del comportamiento de un sistema hamiltoniano caótico. El trabajo de Aref, sumado al hecho de que un flujo bidimensional es mucho más fácil de estudiar en el laboratorio que los flujos tridimensionales, me inspiraron la búsqueda de señales de caos en un sistema experimental de flujo, en un recipiente que mis discípulos y yo construimos, en Amherst, en 1983.

Toma de imágenes

Kenny Leong, un alumno de los cursos superiores, me ayudó a determinar las localizaciones aproximadas de algunos puntos periódicos y de las estructuras de gran escala en flujos bidimensionales mediante un registro cuidadoso de imágenes estroboscópicas del sistema en funcionamiento. (Dado nuestro interés por las mezclas rápidas, nos concentramos en el comportamiento de puntos periódicos de orden bajo, es decir, de períodos uno, dos y tres; los puntos de orden superior no participan en el proceso con tanta frecuencia como lo hacen los puntos de orden bajo.) En un experimento típico, colocamos gotas de un tinte fluorescente en determinadas posiciones del recipiente rectangular, iluminamos el recipiente con luz ultravioleta, hicimos mover sus lados según un determinado esquema y registramos las posiciones y las contorsiones de las gotas tomando fotografías del sistema a intervalos regulares. Si la mezcla es eficiente, las partículas del tinte exploran una gran región del sistema; si la mezcla es pobre, las gotas ceden color sólo lentamente al resto del fluido o permanecen próximas a los puntos periódicos elípticos.

En otros experimentos, Paul D. Swanson, alumno graduado, y yo nos concentramos en flujos que tenían una solución analítica exacta de las ecuaciones del movimiento del fluido. Ello nos facultaba para comparar mejor nuestros resultados experimentales con las predicciones de la teoría. Por desgracia, el número de sistemas para los que disponemos de una solución analítica exacta es bastante exiguo, y muchos se mueven en un terreno ideal sin posibilidad de reproducirlos en los experimentos de laboratorio. Uno de los sistemas que admite solución exacta y se puede repetir experimentalmente es el del flujo entre dos cilindros excéntricos en rotación. Un sistema así ha sido también estudiado por Aref (hoy en la Universidad de California en San Diego) y por Michael Tabor y Rene Chevray, de la Universidad de Columbia.

Los experimentos generalizados so-



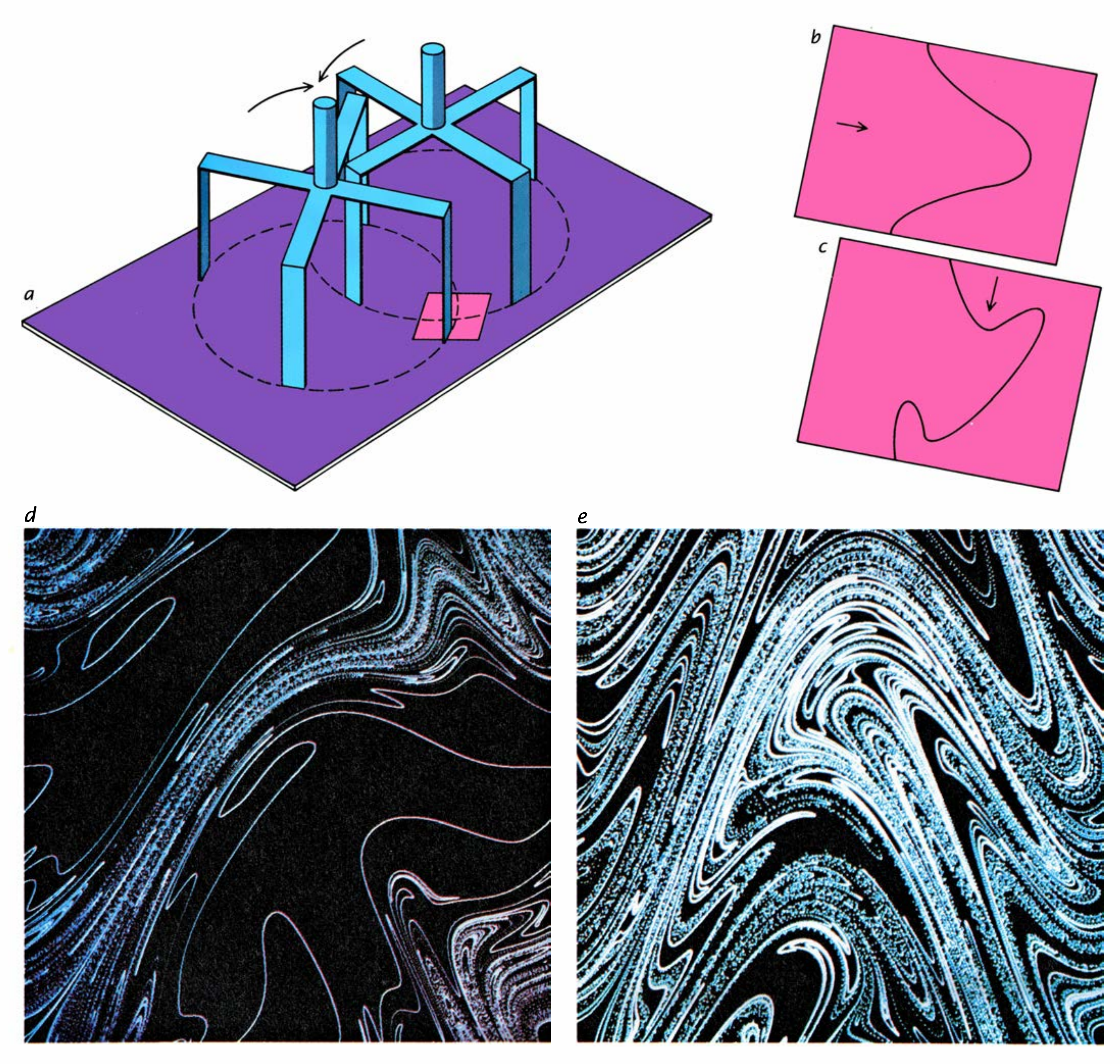
4. PUNTOS ELIPTICOS E HIPERBOLICOS, características típicas de los flujos en dos dimensiones. La fotografía (*abajo*) muestra un flujo así, generado cuando los lados opuestos de un recipiente rectangular lleno de glicerina se mueven en sentidos opuestos a velocidad constante. Las líneas anaranjadas (producidas por un trazador inyectado a lo largo de una línea que se extendía desde la esquina inferior izquierda a la superior derecha del recipiente) están casi alineadas con las líneas de corriente del flujo, las líneas que siguen las partículas del fluido en movimiento en los flujos estacionarios. El esquema del flujo consta de tres puntos fijos: uno hiperbólico central y dos elípticos a cada lado. El flujo cerca de cada punto elíptico (*abajo*) produce una espiral que gira en el sentido de las agujas del reloj; ello aumenta la longitud del trazador linealmente con el tiempo. El flujo en la vecindad de un punto hiperbólico se aproxima al punto en un sentido y se aleja en el otro. Habida cuenta de que la materia fluida no puede cruzar las líneas de corriente, tal flujo estacionario bidimensional no intervendrá en la mezcla. Ahora bien, si se altera el flujo con el tiempo, los hilos estirados del trazador carecen de tiempo suficiente para alinearse con las nuevas líneas de corriente; ello implica que optan entonces por plegarse rápidamente ante un cambio experimentado en la dirección del flujo.

bre flujos caóticos bidimensionales revelan que las macroestructuras de mezcla del fluido (pensemos en las configuraciones y morfologías de las islas y de los grandes pliegues) se pueden reproducir con cierta facilidad; no ocurre lo mismo con los aspectos y pormenores de las estructuras estiradas y plegadas. La razón estriba en que las pequeñas desviaciones en la colocación inicial de las gotas de tinte se agrandan en el interior de las regiones caóticas del flujo. Eso es justo lo que debe suceder: ha de ser imposible reproducir de manera exacta cualquiera de nuestros experimentos de mezcla. Después de todo, el objetivo del proceso de mezcla es crear aleatoriedad. Ello es precisamente lo que se logra mediante los mecanismos de estirar y plegar que se realizan en nuestros experimentos.

Interesa destacar también que el caos puede coexistir con las simetrías del flujo, tales como las representadas por los puntos periódicos. Así, mediante la eliminación sistemática de las simetrías en un flujo caótico, mis colaboradores y yo hemos logrado incrementar el rendimiento del flujo en la mezcla.

Experimentación real y simulación

Si el sistema experimental es sencillo y permite, en consecuencia, deducir una expresión matemática para el campo de velocidades, puede simularse con facilidad en el ordenador. En un programa arquetípico, se colocan estraté-



5. MODELO DE LA BATIDORA desarrollado por John G. Franjione y el autor de este artículo. Sirve para ilustrar el proceso básico de estirar y plegar que tiene lugar en las mezclas (*a*). Una línea dibujada en la superficie de una porción o célula de fluido se estira y se pliega cuando una pala presiona el fluido, primero en una dirección perpendicular a la línea (*b*) y después en una dirección paralela a ella (*c*). La línea se estira sin romperse; cualquier parte que se extienda más allá de la parte de arriba de la célula volverá a entrar por el fondo, y las partes que se extiendan más allá del lado izquierdo volverán a la célula por el lado derecho. Un ordenador puede generar imágenes de la célula que dependen del número de veces que las palas han presionado la célula. En las imágenes que se muestran aquí, se ha estirado y plegado 16 veces una única línea inicial, formada por 100.000 puntos, bajo dos condiciones de mezcla distintas. La mezcla resultante puede quedar confinada a regiones de la célula (*d*) o puede ocupar toda ella (*e*), dependiendo del "vigor" con que las palas presionan el fluido.

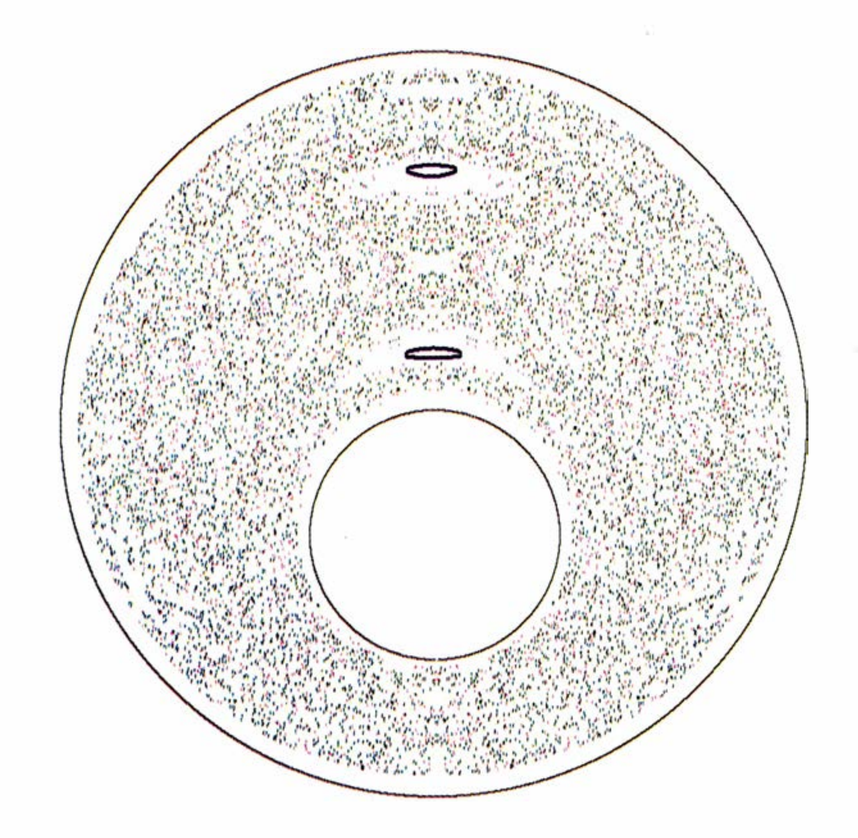
gicamente unas cuantas "partículas" de prueba en un campo de movimientos o de velocidades simulado; los cálculos de las posiciones de las partículas después de unos mil períodos proporcionan una buena imagen del comportamiento del sistema global, transcurrido largo tiempo. La imagen generada por esta clase de simulación se denomina sección de Poincaré. Ahora bien, la sección de Poincaré que muestre un aspecto complejo suele ser prueba "informática" del caos [*véase la figura 6, arriba*]. Las simulaciones por ordenador de las mezclas presentan también cierta forma de irreversibilidad cinemática; en este caso se debe, sin embargo, a la ampliación exponencial de los errores que introduce el ordenador, que sólo puede manejar cifras formadas de un número finito de dígitos.

Si hay simulaciones por ordenador de las mezclas, ¿por qué preocuparse de los experimentos? Lo primero que debe fijarse en la mente es que la resolución de los campos de velocidades precisa ser mucho mayor para simulaciones de mezclas que para simulaciones de la mayoría de los demás problemas de mecánica de fluidos. Incluso campos de velocidades bastante sencillos llegan a producir estructuras de extrema complejidad [*véase la figura 2*]; en algunos problemas de mezclas interesaría resolver algunos de los detalles más finos de las estructuras.

Por ejemplo, en una simulación del flujo del recipiente rectangular, un campo de velocidades calculado al uso puede resultar demasiado basto como para capturar, en sus pormenores, las estrías de los estiramientos y los pliegues. Sería también inútil para señalar las posiciones exactas de los puntos periódicos que determinan el complejo comportamiento de los flujos caóticos. Además, mientras que la mayoría de los problemas de mecánica de fluidos pretende obtener una aproximación del campo de velocidades, en las mezclas el problema no termina con la especificación del campo de velocidades, sino que empieza entonces.

Dificultades

Por esa misma razón, los estudios sobre mezclas se han centrado en meras caricaturas de los flujos (descritas por ecuaciones que en algunos casos pueden resolverse exactamente), y no en problemas reales (cuya solución sólo puede ser aproximada). Lo cierto es que los métodos numéricos para aproximar las soluciones de las ecuaciones





6. FLUJO VISCOSO DE COJINETE: se da entre dos cilindros excéntricos en rotación y puede modelarse mediante ordenador. Si los cilindros se mueven periódicamente en sentidos opuestos, el flujo conduce a una mezcla caótica, tal como se ve en la sección de Poincaré del sistema tras 1000 períodos (*arriba*) y en su esquema de tirones tras 10 períodos (*abajo*). Una sección de Poincaré se genera colocando unas "partículas" de prueba coloreadas en el flujo simulado del sistema que se mezcla, calculando su movimiento para cada período y moviéndolas a sus nuevas posiciones. Un esquema de tirones describe las regiones en las que el fluido se ha estirado en el flujo simulado. La mayor parte del estiramiento tiene lugar dentro de las regiones blancas; el estiramiento es escaso en las regiones de color. El esquema de tirones que se muestra recuerda sorprendentemente la estructura que se produce en el flujo real (*véase la portada de la revista*). Las imágenes fueron realizadas por Paul D. Swanson y el autor, en el laboratorio de la Universidad de Massachusetts en Amherst.

de mecánica de fluidos suelen introducir efectos espúreos que no existen en los problemas de mezcla real.

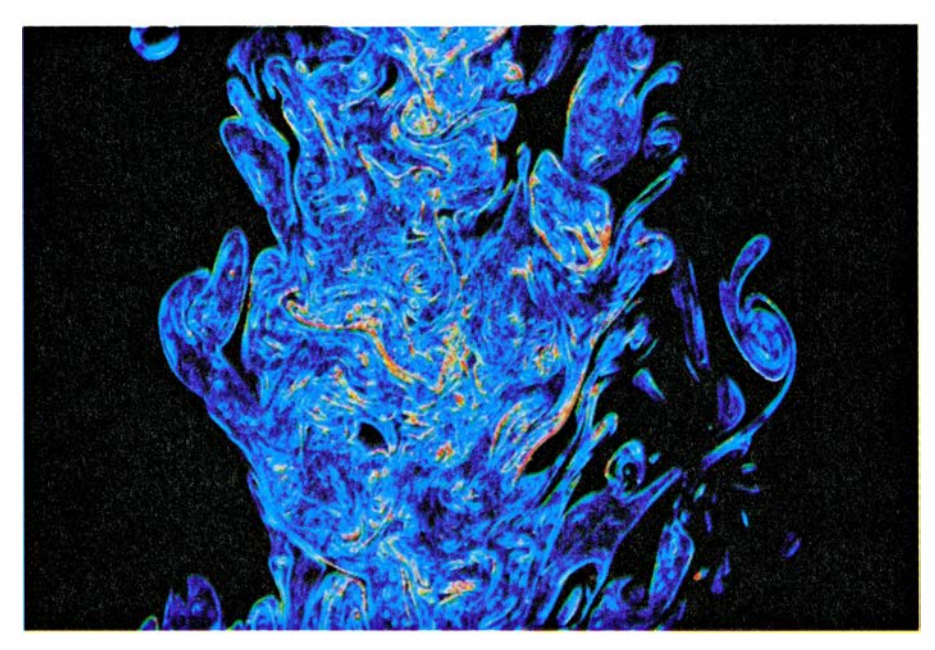
Ni siquiera la reproducción en el ordenador de los flujos simplificados de nuestros experimentos se halla exenta de dificultades insuperables. El computador aborda el fluido como si constara de elementos discretos. En una simulación, cientos de miles de tales elementos pueden constituir una simple gota de tinte, y el número de cálculos involucrados en la representación de su comportamiento caótico en un flujo mezclador puede ser enorme.

Mantener el seguimiento de todas las estrías en regiones de mezcla caótica de un ejemplo relativamente sencillo [*véase el que se muestra en la figura 2*] puede exigir hasta 300 años de tiempo de ordenador en una máquina capaz de realizar un millón de operaciones de coma flotante por segundo. Podría objetarse, ciertamente, que el seguimiento pormenorizado es innecesario, que sería mejor dar cuenta del proceso en un sentido estadístico. Pero, ¿no significaría ello admitir una derrota? Si el campo de velocidades (o de movimiento) puede conocerse exactamente, ¿por qué debe abordarse el problema estadísticamente?

En resumen, los nuevos desarrollos teóricos precisan el concurso de experimentos bien planteados, ya que es bastante probable que la fuerza informática bruta no ayude a contestar muchas de las preguntas relativas a los flujos caóticos. Por ejemplo, ¿qué clases de movimientos deben realizar los lados del recipiente a fin de reducir el tamaño de todas las islas de la cavidad (incluidas las nuevas que puedan aparecer espontáneamente) por debajo de cierto nivel determinado? La contestación a esta pregunta puede permitir el diseño de un sistema refinado de reconocimiento de imágenes que detecte la presencia de islas en una mezcla, y alterar entonces el flujo, a fin de que las islas se mezclen con el resto.

Limitaciones y complicaciones

Antes de poder fabricar máquinas de mezclar "inteligentes", habrá que profundizar mucho más en el conocimiento de los flujos reales. Aunque los experimentos y las simulaciones por ordenador descritas en este artículo proporcionan algunas pistas acerca de los problemas generales de las mezclas —qué hacer para aumentar exponencialmente el área de contacto entre dos fluidos, por ejemplo—, representan situaciones bastante específicas de problemas ideales. Así, los flujos en recipientes descritos en este artículo no tienen inercia. En otras palabras, el flujo se detiene en cuanto las paredes del recipiente dejan de moverse. En consecuencia, tales flujos no conducen a algunos de los procesos que se observan en los flujos turbulentos.



7. FLUJO TURBULENTO, capaz de generar estructuras completamente distintas de las producidas por los flujos lentos y viscosos. La imagen, realizada por K. R. Sreenivasan, de la Universidad de Yale, es una reconstrucción mediante ordenador de un chorro de agua que ha sido expelido por una boca circular sumergida en agua en reposo. Las estructuras del flujo se filmaron inicialmente disolviendo un tinte fluorescente en el agua expelida e iluminando con una lámina de luz láser el eje de la boca. La intensidad de la fluorescencia resultante es proporcional al gradiente de concentración relativa del tinte en el agua; la imagen se coloreó según un código que va desde el azul profundo hasta el rojo, según la intensidad del gradiente de concentración. El flujo turbulento que se desarrolló parece formado por estructuras fractales superpuestas.

Para expresarlo en términos más técnicos, los números de Reynolds (la razón entre las fuerzas de inercia y las fuerzas viscosas en un fluido) de los flujos que hemos estudiado en nuestros experimentos son pequeños. Los flujos caracterizados por números de Reynolds bajos (los llamados flujos laminares) son ordenados y suaves, mientras que los caracterizados por números de Reynolds elevados producen campos de velocidades bastante complicados, que varían con el tiempo y que conducen a una rápida mezcla. Un observador colocado en un punto fijo de nuestro recipiente experimental vería el mismo campo de velocidades sencillo repetirse periódicamente, en vez de observar los campos aperiódicos e impredictibles que se generan en un flujo turbulento. Aunque debemos a la turbulencia que sea más fácil mezclar la leche en el café con una cucharilla (un sistema que tiene un número de Reynolds relativamente alto), que mezclar dos colores de pintura casera con una espátula (un sistema que tiene un número de Reynolds bajo).

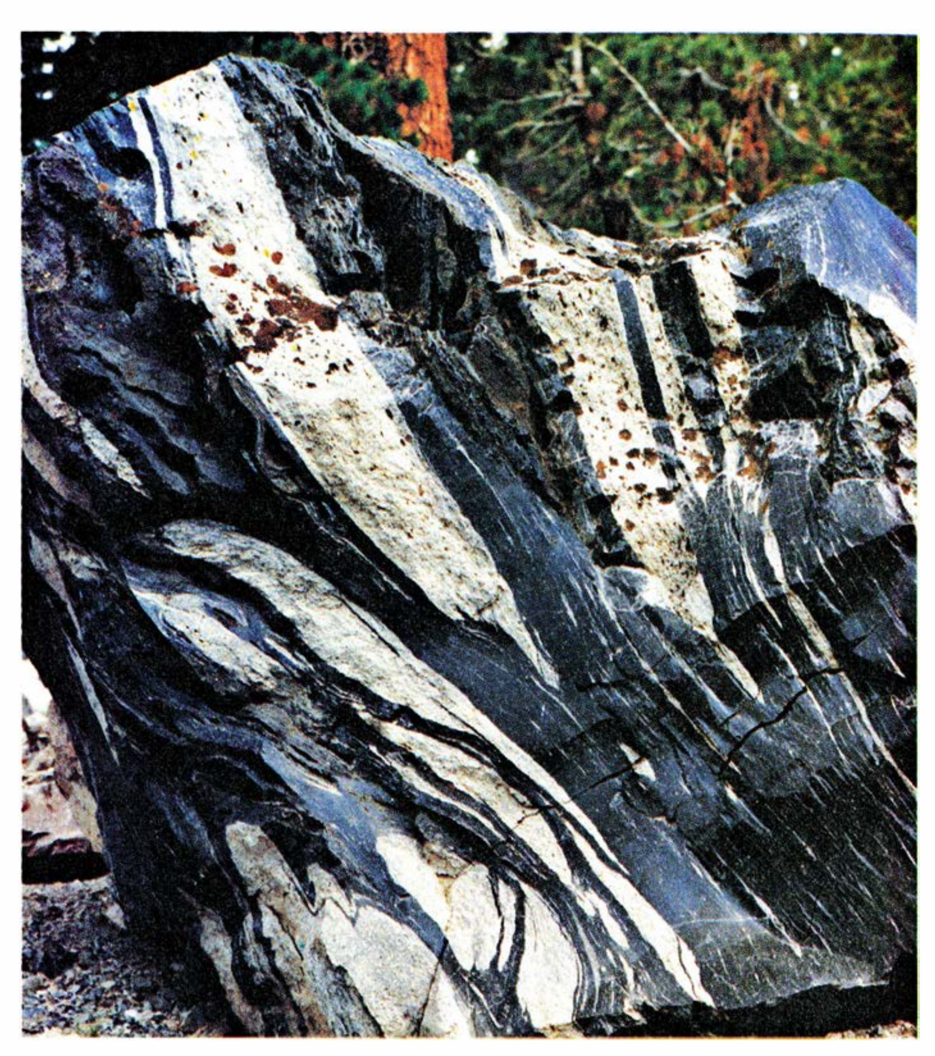
En cierto sentido, hemos excluido, en nuestra exposición, los flujos mezcladores más eficientes (los que son turbulentos); a pesar de ello, hay razones para creer que algunas de las ideas presentadas en este artículo pueden conducirnos a nociones que nos serán de utilidad en el estudio de tales flujos. En esa línea, versiones ligeramente más elaboradas de los flujos caóticos bidimensionales presentan velocidades aperiódicas cuando se miden en un punto fijo. De lo que no cabe duda es que se debe ahondar mucho más si queremos conocer la turbulencia con la precisión con que conocemos los flujos laminares.

A lo largo de la exposición hemos simplificado el asunto al prescindir de la importancia de la difusión en las mezclas. La verdad no es ésa. Para considerar el efecto de la difusión en las mezclas, podemos aplicar un modelo sencillo que establece que la velocidad de difusión entre estrías vecinas de dos materiales miscibles está controlada por la celeridad con que dichas estrías se "comprimen" y se "adelgazan"; esto, a su vez, depende de la componente del flujo según la dirección perpendicular a las estrías. La mezcla ejerce así un doble efecto que potencia la difusión: aumenta el área de contacto entre los fluidos, mientras restringe la distancia en que los fluidos han de difundirse y, al propio tiempo, incrementa los gradientes de concentración. Tal modelo puede generalizarse y explicar

el efecto de la mezcla en la combustión y reacciones químicas similares.

Otro proceso común que —por mor de simplicidad— hemos dejado de lado es el de la rotura en gotitas de fluidos inmiscibles, muy complicada. Existen dos casos límite: un fluido de baja viscosidad, dispersado en otro fluido de elevada viscosidad, y un fluido de alta viscosidad dispersado en otro de viscosidad baja. Ambos casos son difíciles de analizar, aunque por distintas razones. En el primero, el fluido de baja viscosidad se halla sometido al conjunto de fuerzas tangenciales, ya que no puede transmitir con eficacia las tensiones a las gotas del fluido de alta viscosidad. El flujo estacionario tangencial no puede romper la gota cuya viscosidad cuadruplique la del fluido envolvente. Los flujos longitudinales son más potentes, a este respecto, que los flujos tangenciales. Pero también aquéllos podrían mostrar escasa eficacia en el segundo caso (en el que las gotas de baja viscosidad se dispersan en un fluido de alta viscosidad), ya que es necesario comprimir las gotas considerablemente antes de que se rompan.

En Amherst, hemos realizado estudios sobre la mezcla de dos fluidos de distinta viscosidad en nuestro dispositivo experimental. Según cabría esperar, el grado de ruptura es mucho menor dentro de las islas que en las regiones caóticas. Cabría también que una excesiva agitación forzara la unión de las gotitas; de hecho, los fluidos podrían a veces separarse como resultado de su coalescencia. Utilizando sencillos modelos de ordenador, hemos logrado predecir la cinética de tal agregación en flujos caóticos sencillos.



8. LAS ESTRIAS características de las mezclas en fluidos viscosos son patentes en una roca magmática de la cadena volcánica de Inyo, en California oriental. La roca resultó de la mezcla de dos tipos distintos de magmas, uno de los cuales (el que produce las estrías estrechas) contenía pequeñas burbujas de sustancias volátiles. La difusión a través de tales estrías de magma es muy lenta; el tiempo necesario para que la difusión borre estrías del orden de un centímetro de grosor es mayor que la edad de la Tierra. La fotografía fue realizada por Ichiro Sugioka y Bradford Sturtevant, del Instituto de Tecnología de California.

Primera etapa de un largo viaje

Queda por recordar quizá lo más obvio: hasta ahora, todos nuestros experimentos se han desarrollado con flujos bidimensionales. Y el mundo real es un mundo de tres dimensiones. Hace poco, con mis alumnos, construimos el primer aparato capaz de producir experimentos de mezcla controlada en flujos tridimensionales. Hemos empezado a realizar los primeros experimentos. Son muchas las cuestiones fundamentales relativas a las mezclas en flujos lentos tridimensionales, y, por desgracia, parte de cuanto habíamos avanzado en el conocimiento de los flujos bidimensionales no se puede aplicar automáticamente a los tridimensionales.

La lista de los problemas que presenta la mezcla no acaba aquí. La mezcla de fluidos viscoelásticos (fluidos que vuelven a su forma original después de haberse deformado) constituye un problema formidable del que sabemos poco, a pesar de su papel prominente en el procesado de polímeros de elevado peso molecular. La mezcla de fluidos delicados que son incapaces de aguantar elevadas velocidades tangenciales sin degradarse es asunto del mayor interés en bioingeniería. La mezcla de fluidos extremadamente viscosos mediante movimientos térmicos importa a los geofísicos que estudian la mezcla de magmas en el manto terrestre.

A pesar de la abrumadora complejidad de los procesos de mezcla en la naturaleza y en la industria, tenemos la esperanza de que se lleguen a conocer, y esa comprensión se aplicará de manera fructífera en industrias y laboratorios químicos. Puesto que los experimentos sencillos sirven de analogías del caos, pueden, además, poner en claro hechos fundamentales de los sistemas caóticos en general. Los experimentos descritos en este artículo y otros son un primer paso en esa dirección. Hasta ahora, sólo se han explorado algunas de estas ideas; queda todavía largo trecho para la investigación básica y para la explotación técnica.

BIBLIOGRAFIA COMPLEMENTARIA

FLUID MECHANICAL MIXING—LAMELLAR DESCRIPTION. William E. Ranz en *Mixing of Liquids by Mechanical Agitation*, editado por Jaromir J. Ulbrecht y Gary K. Patterson. Gordon y Breach Science Publishers, 1985.

ANALYSIS OF CHAOTIC MIXING IN TWO MODEL SYSTEMS. D. V. Khakhar, H. Rising y J. M. Ottino en *Journal of Fluid Mechanics*, vol. 172, págs. 419-451; noviembre de 1986.

MORPHOLOGICAL STRUCTURES PRODUCED BY MIXING IN CHAOTIC FLOWS. J. M. Ottino, C. W. Leong, H. Rising y P. D. Swanson en *Nature*, vol. 333, n.º 6172, págs. 419-425; 2 de junio de 1988.

Hidratos de carbono y depresión

Algunos trastornos de la conducta identificados durante los últimos diez años se caracterizan por alteraciones del apetito y del comportamiento. Uno de los más conocidos es el trastorno afectivo estacional o TAE

Richard J. Wurtman y Judith J. Wurtman

El 16 de mayo de 1898, el intrépido explorador del Artico, Frederick A. Cook, anotaba en su diario: "El invierno y la oscuridad han ido lenta y progresivamente apoderándose de nosotros... No es difícil leer en las caras de mis compañeros sus pensamientos y su estado de ánimo... La cortina de tinieblas que ha caído sobre el mundo exterior de desolación helada ha descendido también sobre el mundo interior de nuestra alma. Alrededor de la mesa... los hombres se sientan tristes y abatidos, perdidos en sueños de melancolía, de los que, una y otra vez, se despiertan con un rictus estéril de entusiasmo. Por breves momentos, alguien intenta relajar la tensión con algún chiste archisabido. Otros, se esfuerzan por transpirar optimismo; pero todos los empeños encaminados a infundir alguna esperanza fracasan".

Sabemos hoy que los miembros de la expedición de Cook padecían los síntomas clásicos de una depresión invernal, una patología relacionada con una enfermedad psiquiátrica de reciente descripción, denominada trastorno afectivo estacional o TAE. Tal y como se afirma en el apunte citado, el reconocimiento de la asociación entre la depresión y el comienzo del invierno no es algo nuevo. En los últimos años, sin embargo, hemos asistido a un interés creciente por el TAE y dos alteraciones más del comportamiento: la obesidad por ansia de hidratos de carbono (OAHC) y el síndrome premenstrual (SPM), que comparten algunos de sus síntomas. Se incluyen entre éstos la depresión, la letargia y la incapacidad para concentrarse, combinadas con ataques episódicos de sobrealimentación y aumento excesivo de peso. Los síntomas tienden a ser cíclicos, recurriendo en determinados momentos del día (habitualmente a últimas horas de la tarde o durante la noche en la OAHC), del mes (justo antes de la menstruación en el SPM) o del año (durante el otoño o el invierno en el TAE, generalmente).

En el curso de la pasada década se ha venido cosechando abundante información acerca de las manifestaciones clínicas de este grupo de alteraciones de la conducta y del apetito, así como de los procesos bioquímicos que subyacen bajo tales alteraciones. Esas alteraciones vienen instadas por trastornos bioquímicos que se producen en dos sistemas biológicos distintos. En un sistema encontramos la melatonina, hormona que condiciona el talante y la energía subjetiva; en el otro, la serotonina, neurotransmisor que regula la inclinación de una persona hacia los alimentos ricos en hidratos de carbono. La fotoperiodicidad, ciclo diario de luz y oscuridad del planeta, influye sobre ambos sistemas. De hecho, cabría atribuir a la fotoperiodicidad la responsabilidad de los patrones cíclicos de las tres alteraciones.

RICHARD J. WURTMAN y JUDITH J. WURTMAN son, respectivamente, profesor e investigadora científica en el departamento de Ciencias del Cerebro y del Conocimiento del Instituto de Tecnología de Massachusetts. Richard Wurtman, que también es director del Centro de Investigación Clínica del MIT, obtuvo su título de doctor en Harvard. Judith Wurtman entró a trabajar en el MIT en 1974. Ha estudiado zoología y es doctora en biología celular por la Universidad George Washington.

En las latitudes altas de ambos hemisferios, el TAE aparece a finales del otoño o principios del invierno; dura hasta la siguiente primavera. Tras su aparición primera, tiende a recurrir anualmente, a menos que el paciente se traslade a una región donde la duración del día no se acorte significativamente durante el otoño y el invierno. Los enfermos se quejan de ataques episódicos de depresión combinados con un ansia por consumir alimentos ricos en hidratos de carbono. Se acuestan pronto y permanecen en cama de 9 a 10 horas, a diferencia de los enfermos con depresión no estacional, que tienen dificultades para dormir. Sin embargo, su sueño es intermitente y no completamente reparador; durante el día, deambulan soñolientos y les cuesta concentrarse. Cuando llega la primavera, los enfermos de TAE se muestran llenos de energía y creatividad; manifiestan ribetes maníacos en su entusiasmo por la vida. Simultáneamente, su ansia por los hidratos de carbono disminuye y la mayoría de ellos pierden el peso que ganaron durante el invierno.

La historia clínica que traemos a colación ilustra un caso típico de TAE. Enferma de 53 años, maestra, 1 metro 63 centímetros de altura y 82 kilos de peso. Está preocupada por su peso y desde hace años se gasta mucho dinero en dietas efímeras. "Sé que mi problema son los hidratos de carbono: cuando comienzo una dieta no tomo pan, patatas ni dulces y siempre pierdo peso. Pero cuando no estoy a dieta, me pongo nerviosa y tensa hacia media tarde y soy incapaz de concentrarme en lo que estoy haciendo. Necesito comer algo para calmarme, así que compro pasteles o pastas y los voy mordisqueando. Algunas veces, en casa, sigo comiendo hasta que me voy a la cama". Poco después del Día de Acción de Gracias, esta enferma pasó dos meses sintiéndose cansada y deprimida. "Le dije a mi marido que me dejara sola y les puse a mis alumnos ejercicios de problemas para no tener que hablarles en clase. La casa estaba hecha un desastre. No comía nada más que pan y pastas, pero seguía ganando peso. Cuando llegó la primavera empecé a sentirme mejor, quizá porque el año escolar estaba acabando y entrábamos en el verano."

Los síntomas que describe esta enferma son virtualmente idénticos a los

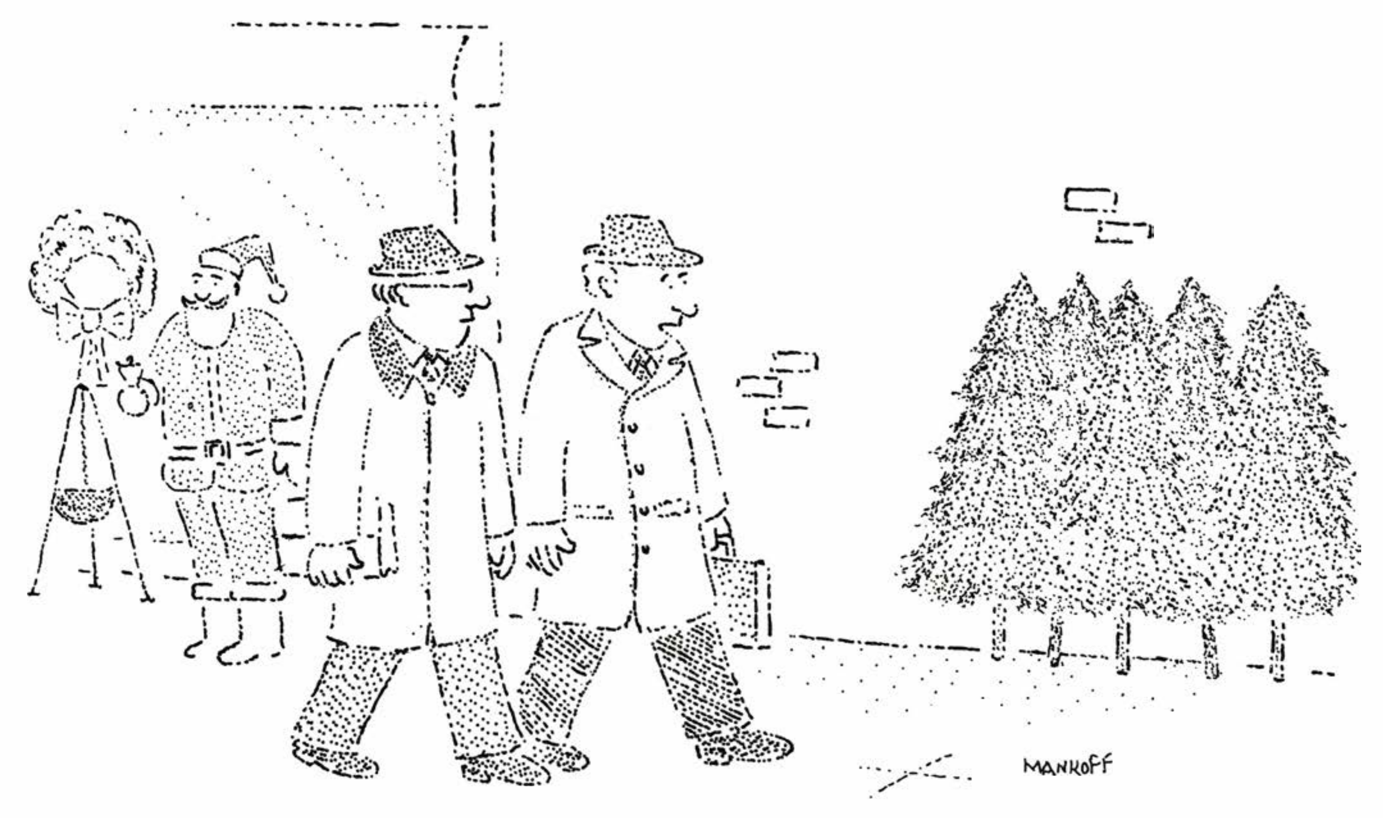
que aparecen en la OAHC y en el SPM, hecha la salvedad de que los que padecen ansiedad por los hidratos de carbono se ven afectados a diario y las mujeres que sufren el SPM manifiestan los síntomas mensualmente, durante la fase lútea del ciclo ovárico, que se mantiene durante las dos semanas previas al inicio de cada período menstrual.

El interés por los trastornos afectivos estacionales surgió a principios de los ochenta, cuando Peter S. Mueller, un psiquiatra del Instituto Nacional de Salud Mental, revisó los datos de una mujer de 29 años que había estado tratando por ataques cíclicos de depresión invernal. En varios años, la paciente estuvo viviendo en distintas ciudades. Mueller mantuvo contacto con ella y observó que, cuanto más al norte vivía, antes manifestaba la depresión durante el otoño y más tiempo persistía deprimida durante la primavera. En dos ocasiones en las que la mujer se trasladó a Jamaica en pleno invierno, la depresión desapareció a los dos días de haber llegado.

Mueller empezó a sospechar que la luz del sol (o la escasez de la misma) contribuía a la depresión de la mujer y decidió experimentar con fototerapia (un tipo de terapia que ya había mostrado su eficacia en el tratamiento de los niños ictéricos y de la psoriasis). Durante varias mañanas consecutivas, expuso a la paciente a 2500 lux de luz, suplementaria, del espectro entero. (Un lux equivale a la iluminación que produce una vela sobre una superficie a un metro de distancia.) En menos de una semana, la paciente se recuperó de su depresión.

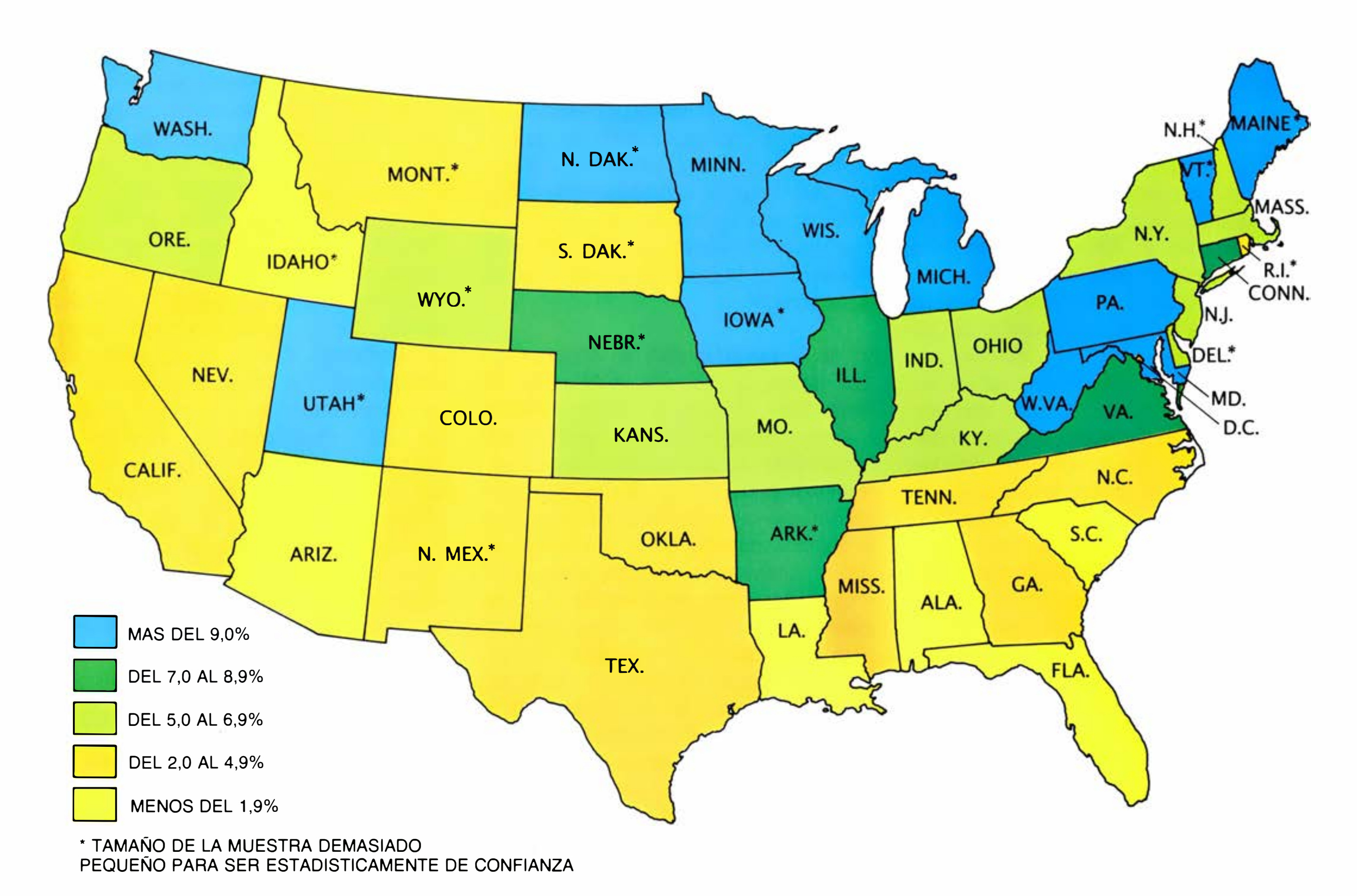
Los descubrimientos de Mueller llamaron la atención a Norman E. Rosenthal, Thomas A. Wehr y Alfred J. Lewy, también del INSM, que estaban interesados en las distintas manifestaciones de la depresión clínica. Emprendieron una investigación exhaustiva de la historia natural de la depresión invernal, reclutando un gran número de voluntarios para observación y tratamiento. Los resultados a que llegaron eran reveladores e intrigantes. Confirmaban el efecto terapéutico de la luz suplementaria en el tratamiento de la depresión invernal mediante fototerapia. Además, sus datos pusieron de manifiesto el nexo de unión entre la depresión invernal y el ansia por los hidratos de carbono.

Un estudio posterior, llevado a cabo por Steven G. Potkin, Daniel F. Kripke, William Bunney y sus colegas, de la Universidad de California en Irvine, proporcionó una información más completa acerca de la correlación entre el TAE y la latitud en los Estados Unidos. Un cuestionario publicado en el periódico *USA Today*, en marzo de 1985, facilitaba una descripción del TAE, pero omitía cualquier referencia a su presumible relación con la duración de la luz del día. Se les pedía a los lectores que respondieran sí o no a 15 preguntas concernientes a la enfermedad. A los que respondían "sí" a 8 o más preguntas (que, posiblemente, padecían un TAE) se les pedía que enviaran el cuestionario a los autores; 723 personas cumplieron el ruego. Se determinó la incidencia de TAE en un estado dividiendo el número de cuestionarios enviados por la venta media diaria del periódico en ese estado. Los resultados indicaban que 100 de cada 100.000 personas de las regiones septentrionales de la Unión padecían TAE; en el sur, la incidencia no llegaba al 6 por 100.000. Tales estimaciones, sin embargo, resultan indudablemente bajas: es menos probable que lean el periódico y contesten a un cuestionario las personas con TAE que las personas que no están afectadas.



"Sí, me siento algo deprimido, pero teniendo en cuenta la estación en que nos encontramos me veo bastante eufórico".

1. MAYOR PREOCUPACION PUBLICA por el trastorno afectivo estacional, o TAE, se viene registrando en los países occidentales a lo largo de los últimos años. El célebre humorista Robert Mankoff captó el espíritu de ese tipo de desorden en este dibujo publicado en *New Yorker* el 10 de diciembre de 1984.



2. INCIDENCIA DEL TAE en los Estados Unidos y su variación con la latitud. En un estado septentrional, como Minnesota, el TAE afecta a más de 100 personas por cada 1000 habitantes, mientras que en Florida afecta a menos de 6 personas por cada 100.000 habitantes. Los asteriscos significan que la muestra es demasiado pequeña para apoyarse en ella. Los datos fueron recogidos por Steven G. Potkin y sus colaboradores, de la Universidad de California en Irvine.

Contemporáneamente, empezamos a estudiar las alteraciones de la alimentación en el Centro de Investigación Clínica del Instituto de Tecnología de Massachusetts, una clínica sin internos en el campus universitario. Un estudio típico en el CIC duraba unas dos semanas. En nuestro caso, valoramos el consumo de hidratos de carbono en 20 pacientes pertenecientes a dos grupos de peso: moderadamente obesos (entre un 20 y un 39 por ciento por encima de su peso ideal) y obesos (entre un 40 y un 80 por ciento por encima de su peso ideal).

Los hábitos alimentarios de nuestro grupo de estudio se registraban cuidadosamente, tanto en las comidas habituales como entre comidas. Las comidas fuera de horario se controlaban mediante un distribuidor automático computarizado (basado en un diseño de J. Trevor Silverstone, del Hospital Clínico Universitario de San Bartolomé, en Londres) que funcionaba las 24 horas del día y contenía una amplia oferta de aperitivos diferentes, desde pastas ricas en hidratos de carbono hasta sardinas ricas en proteínas. Todos los alimentos seleccionados contenían aproximadamente la misma cantidad de grasas (seis gramos, por ejemplo) y de calorías (unas 110). Los alimentos se obtenían tecleando un número de acceso especial en un teclado conectado a un ordenador que mantenía un registro continuo del número y tipo de aperitivos seleccionados por cada paciente. Se pedía a los participantes en el estudio que comieran como solían hacerlo habitualmente y que no se preocuparan por su ingesta calórica; la mayoría cooperaron, confiados en que los datos que obtuviéramos, a la larga, les ayudarían a superar su problema de sobrepeso.

El consumo de alimentos durante las comidas regulares se medía proporcionando a los participantes porciones ilimitadas de alimentos en envases etiquetados y previamente pesados, que se clasificaban según colores y se colocaban en la mesa del comedor. Lo mismo que los aperitivos del distribuidor automático, los alimentos variaban en su contenido en proteínas y en hidratos de carbono y eran equivalentes en su contenido en grasas y en calorías. Tras cada pitanza, el dietista volvía a pesar los envases (y las sobras) para determinar cuánto había comido cada persona de cada alimento.

Nuestros estudios en el CIC nos permitieron cribar y descartar los mitos que rodean a la obesidad, en particular la relación de ésta con el consumo de hidratos de carbono. Creíase que todos los obesos comen demasiado de cualquier alimento apetitoso que puedan consumir en cualquier momento. La verdad es otra: los que sienten ansia por los hidratos de carbono comen solamente demasiados hidratos de carbono y lo hacen en momentos característicos del día. A la hora del almuerzo se parecen a cualquier otra persona, consumiendo un total de 1940 calorías diarias (una mujer adulta media consume un total de 1500 a 2000 calorías y un hombre de 2200 a 2700 calorías). No obstante, hacia el final de la tarde o al principio de la noche los voluntarios empiezan a tomar distintos alimentos y frecuentemente llegan a consumir 800 calorías adicionales o más por persona y día. Se ha observado un

comportamiento similar entre las mujeres con SPM: aumentan su consumo de aperitivos en unas 460 calorías más por día que las mujeres en el mismo período de su ciclo menstrual que no están afectadas por el SPM.

Nos llamó también la atención descubrir (con la ayuda del distribuidor automático) que los pacientes infravaloraban, casi invariablemente, su consumo de alimentos entre comidas. Parece que un aperitivo, si se toma rápidamente, se olvida con facilidad, como si "no contara". Mas, para las personas preocupadas por su peso los aperitivos sí cuentan. En algunos casos, proporcionan hasta un 30 por ciento o más de la ingesta calórica de un individuo.

Descubrimos, por otra parte, que la mayoría de los aperitivos que consumían los enfermos de OAHC o SPM eran de hidratos de carbono. Más de la mitad de las personas obesas, del CIC, con avidez por los hidratos de carbono nunca elegía un aperitivo proteico, aunque la mayoría tomaban proteínas sin problemas durante la comida. Una explicación posible, sin verificar todavía, para esa conducta selectiva con los alimentos sería que la capacidad para regular la ingesta de nutrientes está minada al final de la tarde o al comienzo de la noche en las personas con proclive debilidad por los hidratos de carbono. En una persona sin esa tendencia, el deseo de dulces es poco frecuente, no cíclico y fácilmente satisfecho tomando uno o dos pasteles. Sin embargo, en un sujeto con ansia por los hidratos de carbono el deseo persiste y no se aquieta hasta acabarse nueve o diez pasteles. Ello sugiere que existe una disfunción en el mecanismo de retroalimentación mediante el cual el cerebro sabe que se están consumiendo hidratos de carbono. Otra posibilidad es la de que estos enfermos consumen ese tipo de aperitivos porque los alimentos ricos en hidratos de carbono mejoran su comportamiento, y no porque tengan hambre.

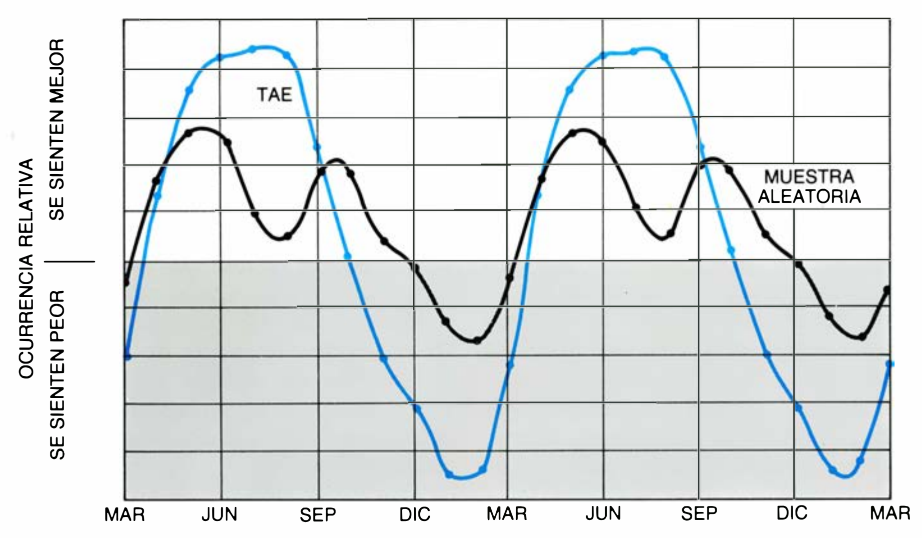
No se sabe a ciencia cierta por qué razón se consumen los aperitivos en determinados momentos del día; esa ocurrencia cíclica, que resulta mensual en el SPM o estacional en el TAE, podría reflejar la acción de las hormonas ováricas o de la melatonina sobre el cerebro, una relación que no ha podido demostrarse en el caso de la OAHC. De cualquier modo, está claro que los aperitivos de hidratos de carbono tienden a exacerbar la obesidad por su elevado contenido en grasas y, por tanto, en calorías.

El apetito inmoderado por los hidratos de carbono constituye, tal parece, una alteración multifacética. Casi dos tercios de los obesos son grandes consumidores de hidratos de carbono, pero no todos los sujetos con debilidad enfermiza por los hidratos de carbono son obesos; muchos controlan su peso haciendo ejercicio, tomando comidas bajas en calorías o satisfaciendo su deseo consumiendo hidratos de carbono bajos en grasa, tales como palomitas de maíz, o gelatina, entre los dulces. Contrariamente, no todos los obesos consumen hidratos de carbono en exceso; los hay sin especial preferencia por los hidratos de carbono, que se sobrealimentan principalmente durante las comidas, pero rara vez ingieren algo fuera del horario habitual.

Nuestra investigación se centró también en los cambios de comportamiento que experimentan los grandes consumidores de hidratos de carbono. Cuando se llevan a cabo pruebas psiquiátricas con esas personas, a través de entrevistas (la escala de Hamilton) o cuestionarios escritos (el inventario de depresión de Beck), se pone de manifiesto su fácil proclividad hacia las depresiones clínicas. Cuando se les pre-



3. MAQUINA DE APERITIVOS del Centro de Investigación Clínica del Instituto de Tecnología de Massachusetts, mediante la cual puede obtenerse información acerca de los alimentos preferidos por los sujetos con y sin inclinación enfermiza por los hidratos de carbono. Ofrece distintos aperitivos con la misma cantidad de grasa y de calorías, pero ricos en hidratos de carbono o en proteínas. Para obtener un aperitivo, la persona introduce un número de acceso a la máquina, que está conectada a un ordenador. De cada sujeto sometido a prueba, quedan registrados los tipos de aperitivos que selecciona y la hora a la que los consume.



4. VARIACIONES ESTACIONALES del comportamiento, frecuentes entre los habitantes de Nueva York (y también en los de otras ciudades nórdicas), pero más graves en los enfermos diagnosticados de TAE. Los datos proceden de un estudio realizado por Michael Terman, del Instituto Psiquiátrico de Nueva York.

gunta por qué sucumben ante los alimentos que saben que favorecen su obesidad, dan una explicación parecida a la que se oye entre las personas que padecen un TAE. Casi nunca está relacionado con el hambre o con el sabor de los alimentos; antes bien, dicen comer para combatir la tensión, la ansiedad o la fatiga mental. Después de comer, se sienten más calmados y con las ideas más claras. Nos preguntábamos si el consumo de cantidades excesivas de hidratos de carbono fuera de las comidas habituales que es causante de obesidad enfermiza no podría equivaler a algún tipo de abuso de sustancias, llevándose a una situación extrema la decisión de consumir hidratos de carbono por sus efectos calmantes y antidepresivos, a pesar de su repercusión negativa sobre la salud y el aspecto físico del sujeto.

Con la ayuda de Harris R. Lieberman y Beverly R. Chew, Judith Wurtman, coautora de este artículo, se dispuso a valorar la relación entre el consumo de hidratos de carbono fuera de las comidas y el comportamiento. Se realizaron las pruebas psicológicas acostumbradas en cuarenta y seis voluntarios, antes y después de tomar una comida rica en hidratos de carbono y sin proteínas; había sujetos con y sin ansia por los hidratos de carbono. Los grandes consumidores de hidratos de carbono se encontraban significativamente menos deprimidos después del aperitivo; los que no presentaban dicha alteración experimentaban cansancio y somnolencia. Estos hallazgos sugieren que los afectados podrían consumir aperitivos ricos en hidratos de carbono para reforzar su vitalidad debilitada, lo mismo que quien toma otra taza de café cuando siente que su energía o su atención están decayendo.

El descubrimiento de que el ansia por consumir hidratos de carbono, al igual que el TAE, presenta una determinada periodicidad, nos llevó a la sospecha de que los fotoperíodos podrían estar relacionados con las manifestaciones cíclicas de las alteraciones del apetito y de la conducta. Sabíamos, por estudios realizados hace 25 años, que la secreción de melatonina seguía un ritmo circadiano específico, asociado con los cambios de luz diarios y estacionales; eso parece corresponderse, al menos conceptualmente, con el ritmo que con mayor frecuencia acontece en el TAE.

En 1958, Aaron B. Lerner y sus colaboradores, de la facultad de medicina de la Universidad de Yale, descubrieron la melatonina, aislándola de glándulas hipofisarias de vacas; observaron, asimismo, que esa sustancia hacía más claros fragmentos de piel arrancados de un renacuajo. Cinco años después, Julius Axelrod y uno de nosotros (Richard Wurtman), por entonces en el INSM, avanzamos la hipótesis sobre la naturaleza hormonal de la melatonina en los mamíferos; nos basábamos en su poder de suprimir la función gonadal cuando se la inyectaba en ratas. Descubrimos luego que la síntesis de melatonina disminuía cuando exponíamos las ratas a la luz, efecto que venía mediatizado por interacciones entre la retina, el cerebro y los nervios del sistema simpático que inervan la hipófisis.

Por la misma época, Wilbur B. Quay, de la Universidad de California en Berkeley, demostraba que los niveles de melatonina en la hipófisis de las ratas presentaban un ritmo diario, aumentando a medianoche y disminuyendo durante el día. Transcurridos algunos años, Russell Pelham y sus colaboradores, de la Universidad de Pittsburgh, describían fluctuaciones similares de la melatonina en el plasma de los seres humanos. Al poco tiempo, uno de nosotros (Richard Wurtman) y Harry J. Lynch, del MIT, descubrimos que los niveles de melatonina en la orina de los humanos presentaban fluctuaciones horarias importantes en muestras obtenidas en los mismos sujetos: la concentración nocturna quintuplicaba, de lejos, los niveles diurnos.

Con el fin de demostrar que el ciclo día-noche, o luz-oscuridad, afecta el ritmo de producción de melatonina, David C. Jimerson, del INSM, Lynch y uno de nosotros (Richard Wurtman) estudiamos los efectos de la inversión brusca de los fotoperíodos de una persona. Reunidos los voluntarios, registramos los niveles de melatonina presentes en su plasma y en su orina; cambiamos luego su fotoperíodo. Los voluntarios se hallaban recluidos en un recinto donde mantuvimos las luces encendidas hasta las 11,00, con lo que el período oscuro diario duraba 12 horas (entre las 11,00 y las 19,00).

Los sujetos necesitaban de cuatro a cinco días para acostumbrarse y ajustarse fisiológicamente al nuevo ciclo de luz secretando melatonina cuando estaba oscuro y suprimiendo la secreción cuando había luz. Quedaba así demostrado que la secreción de melatonina seguía un ritmo circadiano en los seres humanos, como en otros mamíferos, que el ritmo era endógeno (generado por un "reloj" localizado en el cerebro) y que dependía del ciclo luz-oscuridad.

Pero ni otros investigadores ni nosotros logramos demostrar en los humanos lo que Axelrod y Richard Wurtman habían demostrado una década antes en ratas: la secreción de melatonina se suprimía bruscamente si los sujetos se exponían a la luz durante el período oscuro del ciclo. No sin perplejidad, dedujimos que la hipófisis de los humanos era inexplicablemente insensible a los efectos de la luz.

Así las cosas, Lewy descubría, en 1980, que la secreción de melatonina podía suprimirse de manera brusca en los seres humanos por medio de la luz, siempre que ésta fuera suficientemente intensa. Cuando despertaba a los participantes en su estudio a las 2,00 de la

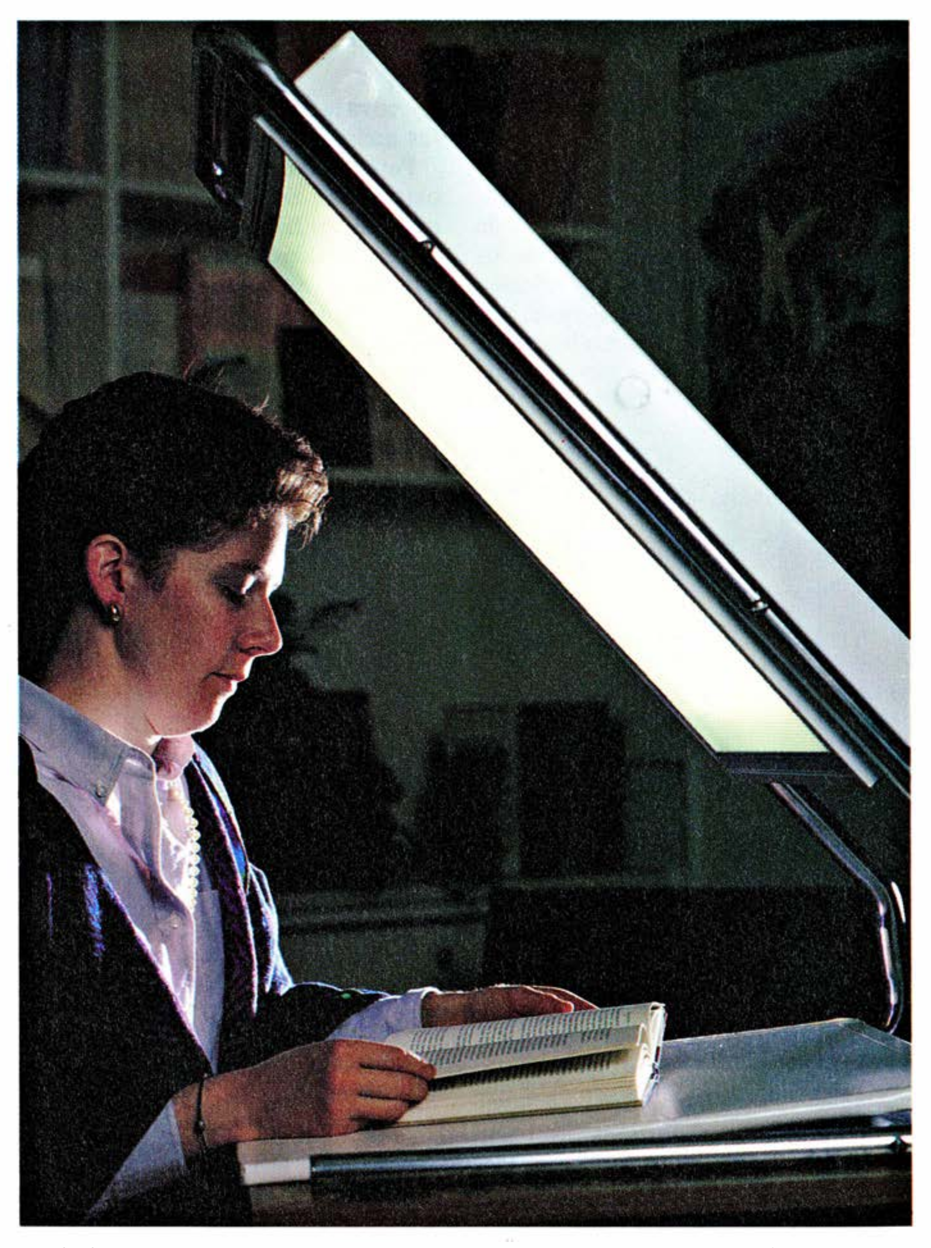
madrugada y les exponía a una luz de 2500 lux durante hora y media, los niveles de melatonina en plasma caían drásticamente. La luz presenta, pues, dos efectos sobre los ritmos de la melatonina en los seres humanos, al igual que en las ratas. Puede alterar el ritmo de la melatonina (como cuando se invertía artificialmente la luz del día en nuestro estudio experimental) o suprimir por completo la producción de melatonina (si se elimina el período oscuro). Una de estas acciones, o ambas, podría ser la responsable del efecto terapeútico de la luz en el TAE.

El trabajo de Mueller, Rosenthal y otros investigadores demostró que, exponiendo los pacientes de TAE a luz suplementaria e intensa durante unas horas cada mañana, desaparecía su depresión y su ansia por los hidratos de carbono a los pocos días. Si bien los obesos consumidores de hidratos de carbono no han recibido todavía tratamiento con fototerapia, hay ya un estudio provisional de Barbara L. Parry, del INSM, que sugiere que la luz suplementaria podría resultar eficaz en el tratamiento de mujeres con SPM, cuyos síntomas empeoran en el invierno.

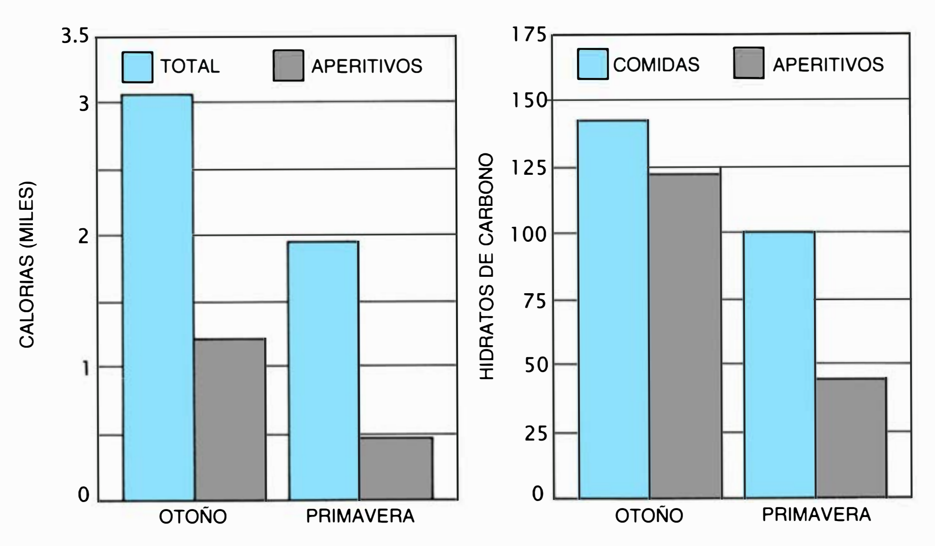
Michael Terman, de la Universidad de Columbia, ha descubierto que exponiendo a los enfermos de TAE a 2500 lux durante dos horas por la mañana se consigue una remisión completa de la depresión y del ansia por los hidratos de carbono en aproximadamente la mitad de los pacientes, generalmente después de tan sólo unos pocos días de tratamiento. La mitad restante manifiesta cierta mejoría, pero no una remisión total. Aunque la investigación no ha concluido, Terman cree en la posibilidad de aumentar la eficacia del tratamiento incrementando el tiempo durante el cual los enfermos se exponen a la luz o intensificando la luz hasta 10.000 lux. La verdad es que tales niveles se aproximarían a la luz solar, que oscila entre 10.000 lux en un día nublado en el norte de Europa y 80.000 lux en un día soleado en el ecuador. Para otros investigadores, sin embargo, lo importante en el tratamiento del TAE sería la duración de la fototerapia, más que su medida. En cualquier caso, se sabe con certeza que la luz debe ser de al menos 2500 lux; la luz habitual en una estancia (cuya intensidad oscila entre 250 y 500 lux) no es capaz de suprimir ni los síntomas del TAE ni la síntesis de melatonina.

La luz administrada por la mañana resulta más eficaz que la administrada a última hora del día. Lewy, Terman y otros han interpretado ese descubrimiento como una prueba de que la luz adelanta el ritmo circadiano de una persona y acorta la fase oscura de la secreción de melatonina. Terman y sus colaboradores han señalado que la disminución de la melatonina en plasma, fenómeno que suele suceder a primera hora de la mañana, se retrasa, unas dos horas, en los enfermos de TAE. Quizá la luz de alta intensidad produce una remisión clínica cuando se administra por la mañana acortando algunas horas el período diario de secreción de melatonina.

¿Es la melatonina la causante del TAE, sea porque se produce en demasiada cantidad o porque se secreta durante demasiado tiempo? ¿O es acaso la melatonina un mero indicador de otro proceso que subyace a la enfermedad? No podemos contestar, de momento, a esas preguntas, pero las pruebas circunstanciales sugieren que existe una relación directa entre la melatonina y el TAE. Lieberman, Lynch y uno de nosotros (Richard Wurtman) descubrimos que la administración de grandes dosis de melatonina a individuos normales inducía somnolencia, debilitaba el estado de alerta y enlentecía el tiempo de reacción. Quizás el inicio de la secreción de melatonina cuando llega la noche sea un importante promotor del sueño, provocando que el cerebro se torne sensible a otros factores inductores del mismo. Ello po-



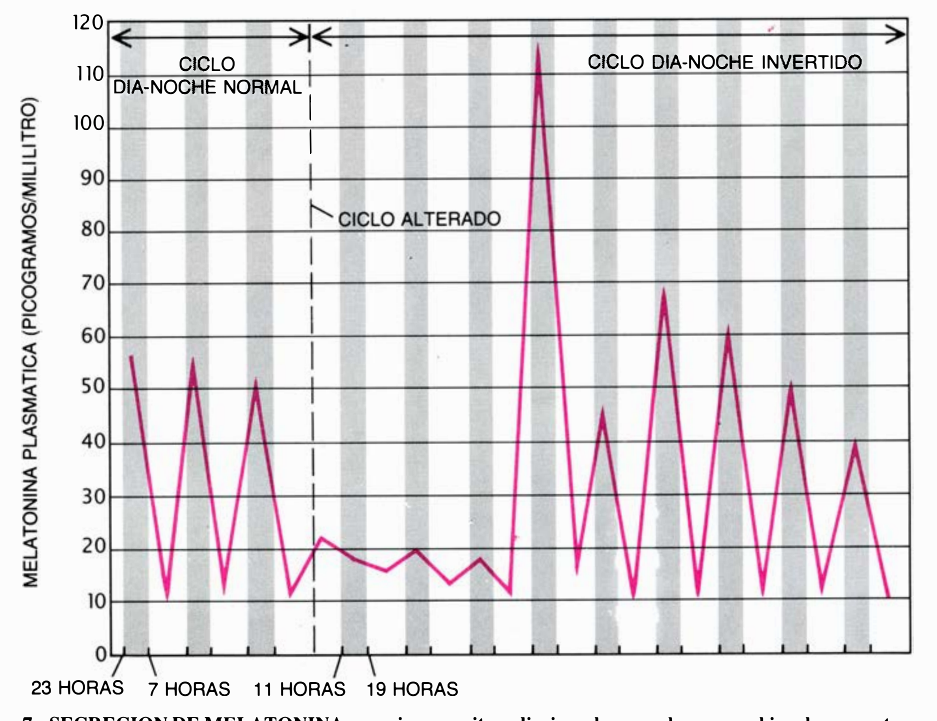
5. FOTOTERAPIA para aliviar la depresión y el ansia por los hidratos de carbono. Los enfermos expuestos por la mañana durante 45 a 60 minutos a luz de alta intensidad mejoran después en dos o tres días.



6. PROPORCION de calorías e hidratos de carbono que consumen los enfermos de TAE en forma de aperitivos; varía según la estación. Durante el otoño, los enfermos toman más de 3000 calorías diarias, de las cuales unas 1200 son en forma de aperitivos; en primavera, su ingesta total de calorías cae por debajo de las 2000 y de ellas menos de 500 son en forma de aperitivos (*izquierda*). En el caso del consumo de hidratos de carbono, aparece un patrón muy similar. En otoño casi el 50 por ciento de los hidratos de carbono consumidos son en forma de aperitivos; en primavera, la proporción baja hasta el 30 por ciento (*histograma de la derecha*).

dría explicar por qué los enfermos de TAE son hipersómnicos en invierno, cuando el período de oscuridad diario es casi el doble que en el verano. El hecho de que la administración oral de melatonina agrave la depresión de los enfermos sugiere también una relación entre la melatonina y la conducta; por desgracia, no se ha desarrollado ninguna droga capaz de bloquear de forma selectiva la producción de melatonina o sus acciones.

Pero, ¿por qué los pacientes con TAE, OAHC o SPM tienen tendencia a consumir muchos hidratos de carbono fuera de las comidas? ¿Por qué sólo determinadas personas son vulnerables a la OAHC? ¿Y cómo, normalmente, sabe el cerebro cuándo se han consumido o se deben consumir hidratos de carbono?



7. SECRECION DE MELATONINA, que sigue un ritmo diario en los seres humanos al igual que en otros mamíferos. Durante el día (*columnas en blanco*), se suprime la secreción de la hormona en la glándula hipofisaria; los niveles de melatonina en plasma son bajos. Durante la noche (*columnas oscuras*), la hipófisis libera melatonina y suben sus niveles en plasma. Si se invierte bruscamente el ciclo de luz diaria en 12 horas, de forma que el período de oscuridad se dé entre las 11 y las 19 horas (en lugar de entre las 23 y las 7 horas), transcurren algunos días antes de que se instaure (se ajuste) el nuevo ciclo de secreción de melatonina en el nuevo fotoperíodo. Una vez ajustada, la secreción de melatonina sigue de nuevo un patrón rítmico.

Los habitantes de los países desarrollados ingieren del 12 al 14 por ciento de las calorías en forma de proteínas y unas tres o cuatro veces más en forma de hidratos de carbono. Hasta el oso renuncia ocasionalmente a la miel por un pescado. ¿Cómo se regula ese fenómeno? Sabemos ya que en la respuesta a esas preguntas interviene la serotonina, un neurotransmisor o sustancia liberada por una neurona cuando se excita y que transmite el impulso nervioso a través de la sinapsis hasta la neurona siguiente.

La serotonina es un derivado del triptófano, un aminoácido que se encuentra presente habitualmente a bajas concentraciones en la sangre. La proporción de hidratos de carbono en la dieta de una persona afecta la tasa de conversión: los hidratos de carbono estimulan la secreción de insulina, que facilita el suministro de aminoácidos a los tejidos periféricos, como el músculo. Ahora bien, los niveles de triptófano en sangre no se ven afectados por la insulina, y así la proporción relativa de triptófano en sangre con respecto a otros aminoácidos aumenta cuando se consumen hidratos de carbono. Por competir el triptófano con otros aminoácidos por el transporte a través de la barrera hematoencefálica, la secreción de insulina acelera su entrada en el sistema nervioso central, donde se introduce, entre otras células, en una familia especial de neuronas denominadas núcleos del rafe. Allí, se convierte en serotonina.

A su vez, el nivel de serotonina interviene en un mecanismo de retroalimentación que influye sobre la cantidad de hidratos de carbono que un individuo decide consumir [véase "Alimentos que modifican la función cerebral", de Richard J. Wurtman; INVESTIGACIÓN Y CIENCIA, junio de 1982]. Cuando se altera el mecanismo de retroalimantación, tal y como pensamos que sucede cíclicamente en los enfermos con TAE, OAHC o SPM, el cerebro no puede saber cuándo se consumen hidratos de carbono y, en consecuencia, el deseo por tomarlos persiste más tiempo del que debe.

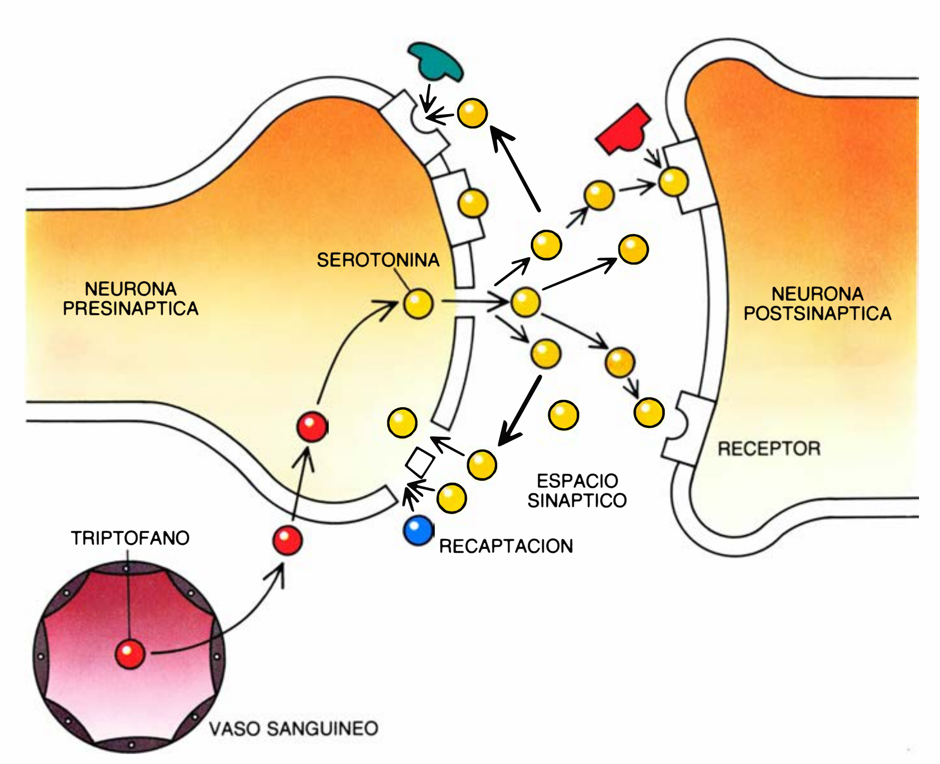
La serotonina regula también otras conductas, incluido el estado de ánimo y el sueño. Bonnie Spring, actualmente en la facultad de medicina de la Universidad de las Ciencias de la Salud de Chicago, descubrió que mujeres sin especial ansia por los hidratos de carbono se mostraban somnolientas y cometían errores después de consumir una co-

mida muy rica en hidratos de carbono (que supuestamente aumenta los niveles de serotonina en el cerebro). Lieberman y uno de nosotros (Judith Wurtman) observamos respuestas similares entre individuos obesos sin proclividad enfermiza hacia los hidratos de carbono. Talante que contrasta con el de los grandes consumidores de hidratos de carbono; éstos se sienten frescos y revigorizados después de una comida rica en hidratos de carbono.

Los mecanismos que regulan las proporciones relativas de hidratos de carbono y proteínas que se consumen en la dieta son más evidentes cuando el círculo de retroalimentación se rompe, como sucede en un enfermo que está tomando fármacos que modifican la neurotransmisión mediada por la serotonina. Si se permite que las ratas de un experimento elijan entre dos o más comidas sintéticas que contengan diferentes proporciones de hidratos de carbono y proteínas, normalmente alternarán entre ellas. Pero si se les inyecta serotonina (en el cerebro) o se les administran fármacos que potencien los efectos de la serotonina facilitando su liberación en las sinapsis nerviosas, prolongando su actividad o estimulando sus receptores, las ratas reducirán selectivamente su ingesta de hidratos de carbono.

Los ensayos de fármacos llevados a cabo con seres humanos demuestran que una droga potenciadora de la serotonina, la *d*-fenfluramina (que libera serotonina en las sinapsis del cerebro y prolonga después su acción bloqueando su recaptación en la neurona presináptica), presenta un efecto similar, suprimiendo selectivamente el consumo de hidratos de carbono fuera de las comidas en pacientes de obesidad por ansia de hidratos de carbono (OAHC).

Asimismo, hemos descubierto, en colaboración con Dermot A. O'Rourke, psiquiatra del Hospital General de Massachusetts, que la *d*-fenfluramina resulta eficaz en el tratamiento del TAE; limita el consumo de hidratos de carbono (y el aumento de peso asociado al mismo) y mejora simultáneamente los síntomas de depresión. Más recientemente, con Amnon Brzezinski, de la facultad de medicina Hadassah de la Universidad Hebrea de Jerusalén, descubrimos que la *d*-fenfluramina colaboraba en el tratamiento de los síntomas equivalentes en las pacientes con SPM. En 12 de 17 mujeres estudiadas, la admisnistración de la droga durante un período de unos seis meses aquietó el anhelo enfermizo de hidratos de carbono y frenó la depresión.



8. **SE ENCARGA LA SEROTONINA de regular el consumo de hidratos de carbono. El proceso comienza con el aminoácido triptófano (*naranja*), que circula por sangre hasta el cerebro, donde se localiza en núcleos del rafe. Después de introducirse en la neurona presináptica, el triptófano se convierte en serotonina siguiendo un proceso de dos pasos (*amarillo*). Entonces la serotonina se libera al surco o hendidura sináptica que separa, de la postsináptica, la neurona presináptica. La serotonina que alcanza la neurona postsináptica se une a unos receptores especiales. Los niveles de serotonina aumentan en respuesta al consumo de hidratos de carbono. Conforme se libera mayor cantidad de serotonina, se transfiere más información a la neurona postsináptica, donde se activa un mecanismo de retroalimentación. Cuando su concentración es alta, la serotonina se une a los receptores presinápticos, suprimiéndose la liberación de serotonina adicional por parte de la neurona presináptica. También puede retirarse de la sinapsis, siendo recaptada en la neurona presináptica. Los fármacos que estimulan la liberación de serotonina (*verde*) o las drogas que bloquean su recaptación (*azul*) aumentan la transmisión de información a través de la sinapsis y disminuyen el consumo de hidratos de carbono fuera de las comidas; los fármacos que bloquean los receptores serotonínicos (*rojo*) aumentan el apetito, particularmente por los hidratos de carbono. (Ilustración realizada por A. Christie.)**

Otra alteración, que suponemos podría estar relacionada con la serotonina (y, por tanto, con el TAE, la OAHC y el SPM), es un tipo de bulimia asociada con consumo elevado de alcohol, frecuentemente de bebidas ricas en hidratos de carbono, pero con pocos vómitos o ausencia de los mismos. La mayoría de tales pacientes son mujeres medianamente obesas; muchas de ellas padecen depresiones severas y provienen de familias con historia de depresión y de abuso de alcohol. G. F. M. Russell, de la Universidad de Londres, y Arthur G. A. Blouin, de la Universidad de Ottawa, realizaron ciertos estudios preliminares que sugerían que la *d*-fenfluramina podía ser eficaz en el tratamiento de tales mujeres; las que respondían a los efectos antidepresivos de la droga tenían más probabilidad de beneficiarse también de sus efectos supresores del apetito.

Hemos descubierto que, por regla general, las drogas que facilitan selectivamente la neurotransmisión mediatizada por la serotonina (así la *d*-fenfluramina, la femoxetina, la fluoxetina, la zymelidina y la fluvoxamina) tienden a inducir una pérdida de peso, mientras que las drogas que bloquean la transmisión mediatizada por la serotonina o los antidepresivos que interaccionan con otros neurotransmisores distintos de la serotonina presentan el efecto opuesto: despiertan el deseo de hidratos de carbono y, consecuentemente, aumento de peso.

Nadie podría afirmar razonablemente que los síntomas del TAE, la OAHC o el SPM carecen de trascendencia. Los períodos prolongados de depresión profunda y de irritabilidad pueden comprometer gravemente la capacidad de una persona para mantener su nivel de relación personal. Pero tampoco cae fuera de la normalidad sentirse bajo de ánimo, abusar en ocasiones del chocolate o las pastas, ganar kilos en invierno o cambiar de estado de ánimo cuando llega la menstruación.

De hecho, los cambios estacionales en la conducta afectan tanto a las personas normales como a las que padecen TAE. De 200 sujetos elegidos aleatoriamente en la guía de teléfonos de Nueva York y encuestados por Terman y sus colaboradores, la mitad declaraban perder parte de sus energías en otoño y en invierno. Un 47 por ciento afirmaban que ganaban peso en ese período, un 31 por ciento que dormían más y otro 31 por ciento que perdían interés en las actividades sociales. Entre los encuestados que declaraban cierta disminución de sus energías en determinada época del año, aproximadamente el 50 por ciento lo referían al otoño y al invierno; sólo el 12 por ciento se sentían así en verano. Terman concluyó que un porcentaje significativo de la población de Nueva York sufre un TAE leve; sospechamos que los habitantes de otras ciudades nórdicas presentan una incidencia similar.

En Tromø, Noruega, a 69 grados de latitud el sol no se eleva del horizonte entre el 20 de noviembre y el 20 de enero; se cifra en un 24 por ciento la población afectada por insomnio invernal. Charles S. Mullin, de la Academia Naval de los Estados Unidos, ha descrito una incidencia importante de falta de sueño, depresión, irritabilidad, perturbación de la capacidad intelectual y aumento de peso de hasta 9 a 14 kilos entre los científicos y el personal militar que pasan el invierno en la Antártida. El hecho de que el TAE alcance su máxima incidencia en el hemisferio sur durante los meses de junio y julio indica que no se trata simplemente de una forma de tristeza por las vacaciones o el resultado de reflexiones melancólicas causadas por el paso de otro año más.

¿Es, pues, el TAE una exageración de una respuesta normal de los seres humanos a los niveles menores de luz durante el otoño y el invierno? ¿Es quizá similar a la hibernación? Probablemente no. Los animales que hibernan rebajan la temperatura de su cuerpo, suspenden la actividad reproductora y pasan el invierno en un letargo profundo. Las personas con TAE no realizan ninguna de esas acciones. Ni siquiera limitan el tiempo que pasan en sueño profundo (medido a través de un electroencefalograma).

El estilo de vida que lleva la sociedad actual aumenta la vulnerabilidad a la depresión estacional al recortarse el tiempo que nos hallamos expuestos a la luz: Daniel Kripke y sus colaboradores

midieron la cantidad de tiempo por día que ancianos sanos de San Diego —una región con un clima particularmente favorable— se encontraban expuestos a la luz solar. Para nuesta sorpresa, los hombres estaban al sol sólo 75 minutos por cada período de 24 horas y, 20 minutos, las mujeres. No todos necesitamos vivir en California o en Marbella, pero quizá la mayoría de las personas precisa verse expuesta a más luz, algo que era natural en nuestros antepasados. Lo mismo que un oficinista necesita acudir al gimnasio para compensar la falta de ejercicio, las personas que trabajan en lugares cerrados necesitan exponerse a la luz más adecuadamente.

Queda mucho por investigar sobre la conducta y las alteraciones del apetito y sobre la relación entre la serotonina y la melatonina. ¿Por qué un enfermo con TAE responde bien a la luz suplementaria, que, presumiblemente, actúa influyendo sobre la melatonina, y a las drogas que estimulan la liberación de la serotonina? ¿En qué momento de la secuencia de los procesos patofisiológicos que conducen a la aparición de un TAE actúan tales tratamientos? Antes de que podamos contestar esas dos preguntas, interesaría saber si la luz o la melatonina tienen un efecto directo sobre las neuronas que liberan la serotonina. Mientras no dispongamos de mejores respuestas, habrá que conformarse con las nuevas terapias que al menos se muestran eficaces contra esas alteraciones, ya que no podemos todavía saber qué mecanismos operan bajo esos tratamientos.

BIBLIOGRAFIA COMPLEMENTARIA

SEASONAL AFFECTIVE DISORDER: A DESCRIPTION OF THE SYNDROME AND PRELIMINARY FINDINGS WITH LIGHT THERAPY. Norman E. Rosenthal, David A. Sack, Christian Gillin, Alfred J Lewy, Frederic K. Goodwin, Yolande Davenport, Peter S. Mueller, David A. Newsome y Thomas A. Wehr en *Archives of General Psychiatry*, vol. 41, n.º 1, págs. 72-80; enero de 1984.

D-FENFLURAMINE SELECTIVELY SUPPRESSES CARBOHYDRATE SNACKING BY OBESE SUBJECTS. Judith Wurtman, Richard Wurtman, Sharon Mark, Rita Tsay, William Gilbert y John Growdon en *The International Journal of Eating Disorders*, vol. 4, n.º 1, págs. 89-99; febrero de 1985.

ON THE QUESTION OF MECHANISM IN PHOTOTHERAPY FOR SEASONAL AFFECTIVE DISORDER: CONSIDERATIONS OF CLINICAL EFFICACY AND EPIDEMIOLOGY. Michael Terman en *Journal of Biological Rhythms*, vol. 3, n.º 2, págs. 155-172; 1988.

La caza de *Proconsul*

Sesenta años después de su descubrimiento, Proconsul *ha dejado de ser considerado el antecesor del chimpancé y el gorila, para convertirse en el último de los antepasados comunes a los grandes simios y al hombre*

Alan Walker y Mark Teaford

El simio prehistórico *Proconsul* es, en la actualidad, el mejor conocido de todos nuestros antepasados, pero su trayectoria, desde la oscuridad de una zanja hasta la fama en los círculos científicos, está tan jalonada de giros y sorpresas que constituiría un buen guión para un serial radiofónico. Es la suya una historia de despropósitos. Varios fragmentos de importantes ejemplares, una vez desenterrados, se separaron y se enviaron a centros muy distantes entre sí, hasta que la fortuna volvió a reunirlos décadas más tarde. Es también una historia con final feliz, ya que las campañas realizadas en los últimos tiempos han logrado recolectar cerca de 800 nuevos ejemplares de primates hominoideos (la superfamilia de primates que incluye los grandes simios, los simios inferiores y los seres humanos). Estos nuevos hallazgos han enriquecido enormemente el muestrario de fósiles de *Proconsul*, al tiempo que erigen este género en modelo idóneo de lo que fue el último antepasado común de los grandes simios y del hombre.

La historia comienza en 1927, en el momento en que H. L. Gordon, un colono establecido en la parte occidental de Kenya, encontró restos fósiles cuando extraía calizas de una cantera. Pensando que podrían encerrar algún interés, Gordon los envió al paleontólogo A. Tindell Hopwood, del Museo Británico. A primera vista, uno de los fósiles no era más que un simple diente que sobresalía desde el interior de un nódulo de roca del tamaño de la palma de la mano. Sin embargo, cuando se eliminó la matriz rocosa, resultó que el ejemplar correspondía al maxilar izquierdo, o mandíbula superior izquierda, de un primate hominoideo. Los restantes fósiles que se encontraron en el yacimiento indicaban que el depósito pertenecía al Mioceno Inferior, hace unos 18 millones de años.

Por aquel entonces los fósiles de simios eran mal conocidos y aún no se había descubierto ninguno con semejante antigüedad; de ahí el extraordinario interés de aquellos ejemplares. Con todo, Hopwood decidió no hacer público el hallazgo hasta disponer de nuevos datos que confirmasen que lo que había encontrado correspondía a una nueva forma de primates. Después de recabar fondos para una campaña, en 1931 se trasladó a Kenya, donde le acompañaría el éxito en su búsqueda de material adicional de hominoideos fósiles. Dos años más tarde, publicó sus hallazgos, manifestando su convencimiento de que la mandíbula encontrada por Gordon era la de un nuevo género antepasado del chimpancé.

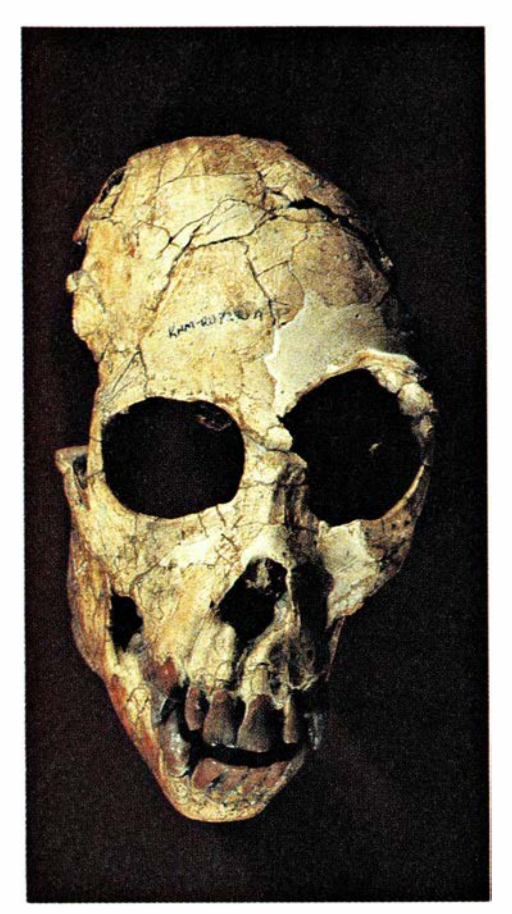
Los londinenses aficionados a los espectáculos de variedades se divertían aquellos días con el número de un chimpancé trajeado, calada la gorra, que montaba en bicicleta y fumaba en pipa. Le llamaban "Cónsul"; en un rasgo de humor a Hopwood no se le ocurrió otra cosa que denominar *Proconsul africanus* al nuevo antropoide.

Las páginas siguientes de esta historia llevan la firma de Louis y Mary Leakey. En los años cuarenta y comienzos de los cincuenta realizaron varias campañas en el occidente de Kenya. En 1948 Mary Leakey halló, en la isla Rusinga del lago Victoria, un cráneo que llegaría a ser el ejemplar más famoso de *Proconsul*. Cuando ella lo encontró (en una ladera de materiales sedimentarios blandos), la zona posterior del cráneo estaba expuesta a la intemperie, por cuya causa se había erosionado parcialmente. La parte frontal y la mandíbula estaban casi enteras; pero la zona posterior y las laterales aparecían representadas por sólo algunos fragmentos recuperados en las partes más bajas de la pendiente.

El cráneo encontrado en 1948 se atribuyó a la misma especie a la que pertenecía la mandíbula hallada por Gordon y descrita por Hopwood. Sin embargo, otros descubrimientos realizados por los Leakey sugerían que en la isla Rusinga y en la de Mfango, vecina a la anterior, podrían haber vivido dos especies distintas de *Proconsul*: una especie grande, de la talla del chimpancé, a la que se denominó *Proconsul nyanzae*, y otra de menor tamaño, *Proconsul africanus*, representada por la mandíbula de Hopwood y el cráneo hallado en 1948.

El siguiente ejemplar importante fue descubierto en 1951 por el geólogo Tom Whitworth, cuando investigaba el área de Kiakanga en la isla Rusinga. En una chimenea vertical de rocas verdosas que atraviesa unos limos grises sin fósiles, Whitworth halló un esqueleto de cerdo y diversos huesos más. Entre los huesos que pudieron ser extraídos de la durísima matriz en la que estaban incluidos, aparecieron piezas de un cráneo perteneciente a un individuo subadulto (juvenil tardío), gran parte de un miembro anterior, una mano y fragmentos de un pie, todo ello correspondiente a un mismo ejemplar de *Proconsul*.

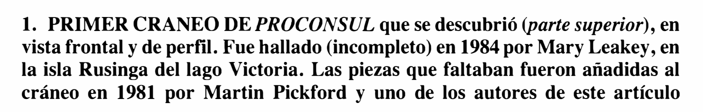
ALAN WALKER Y MARK TEAFORD han trabajado juntos desde 1984 en la búsqueda de fósiles de simios del Mioceno Inferior de Kenya. Son, respectivamente, profesor y profesor ayudante del departamento de biología celular y anatomía de la facultad de medicina de la Universidad Johns Hopkins. Walker se doctoró en paleontología y anatomía por la Universidad de Londres. Colabora con Richard Leakey y otros colegas del Museo Nacional de Kenya en el estudio de la evolución de los primates. Teaford ha alcanzado su doctorado en antropología por la Universidad de Illinois en Urbana-Champaign en 1981. Ha centrado su trabajo en el estudio del uso y función de los dientes de los primates y en la dieta de animales extintos. En estas investigaciones, ha hecho frecuente utilización del microscopio electrónico.











1. PRIMER CRANEO DE *PROCONSUL* que se descubrió (*parte superior*), en vista frontal y de perfil. Fue hallado (incompleto) en 1984 por Mary Leakey, en la isla Rusinga del lago Victoria. Las piezas que faltaban fueron añadidas al cráneo en 1981 por Martin Pickford y uno de los autores de este artículo (Walker), quienes las habían encontrado mezcladas con huesos de tortuga. Los huesos oscuros de esta mano de *Proconsul* (*abajo a la izquierda*) y del pie (*derecha*), fueron descubiertos en 1951 por el equipo de Louis Leakey. Los huesos claros se recuperaron 30 años más tarde en el Museo Nacional de Kenya.

Creíase por entonces que la chimenea de roca verde, compuesta por la acumulación de cenizas volcánicas gruesas transportadas por agua, correspondía al relleno de una especie de sima, excavada por un río, a la que habrían ido a parar los restos de los animales y en donde más tarde fosilizarían. Si ese era el caso, los investigadores deberían enfrentarse a dos problemas: origen de la mezcolanza faunística y edad respectiva de fósiles y medio. Puesto que esqueletos enteros y huesos sueltos podían haber sido arrastrados desde cualquier punto situado aguas arriba, cabía la posibilidad de que la asociación de fauna presente en la chimenea difiriese bastante de la comunidad real en la que había vivido *Proconsul*. Considerando, en segundo lugar, que el agua pudo abrir el boquete en los sedimentos preexistentes, ya compactados, el relleno del mismo pudo producirse en época muy posterior a aquella en la que tuvo lugar el depósito de los materiales circundantes, cuya edad hemos cifrado en 18 millones de años.

Recientemente, han vuelto a entrar en escena los hallazgos de 1948 y 1951 para representar nuevos papeles en la historia de *Proconsul*. Este "renacimiento" comenzó hace escasos años cuando Martin Pickford, del Instituto de Paleontología de París, reparó en una anotación realizada por Louis Leakey en el cuaderno de campo que este último había utilizado en 1947. La nota se refería a la presencia de huesos de cráneo, posiblemente pertenecientes a primates, que se habían recogido en el mismo lugar donde Mary Leakey encontraría el cráneo un año más tarde. Pickford comprendió rápidamente que estos fragmentos podían corresponder a las partes del cráneo desprendidas por la erosión. Se las arregló para averiguar el paradero de las piezas, hasta que, finalmente, las encontró en una colección de huesos de tortuga depositada en el Museo Nacional de Nairobi. Eran, sin duda alguna, las piezas perdidas de *Proconsul*, pertenecientes a la parte posterior del cráneo.

Pickford y uno de nosotros (Walker) nos las ingeniamos para pegarlas de nuevo al original, logrando así un cráneo completo, desde el hocico hasta la parte superior y hasta el foramen magno (el amplio orificio que se sitúa en la base del cráneo).

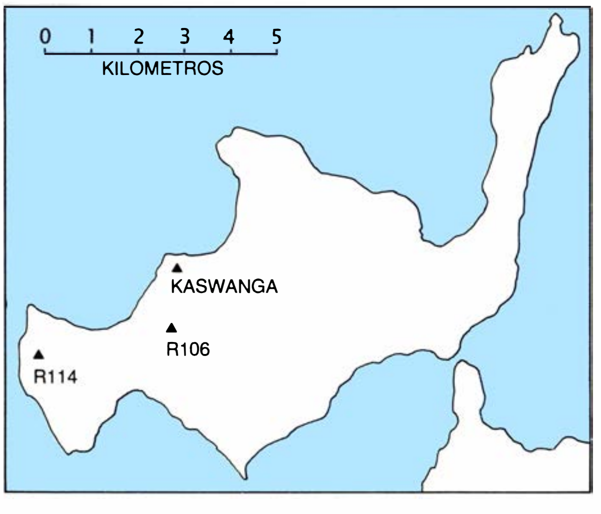
Con un cráneo más completo, esperábamos ya disponer de una fuente de información de suma importancia para los biólogos evolutivos: conocer cuán encefalizado (o "cerebrado") se hallaba *Proconsul*. El grado de encefalización viene dado por la relación entre el volumen del cerebro y el peso del cuerpo; a este respecto, el cráneo de *Proconsul* nos permitía efectuar una buena estimación de la capacidad craneal de un hominoideo del Mioceno. Para estudiar la capacidad craneal se suele medir el volumen de agua desplazado por un molde interno del cráneo. Como es obvio, tal procedimiento requiere disponer de un cráneo no deformado; por desgracia, el ejemplar de *Proconsul* estaba algo aplastado y doblado.

Sin embargo, Pickford, Dean Falk, de la Universidad estatal de Nueva York en Albany, Richard J. Smith, de la Universidad de Washington, y uno de nosotros (Walker) idearon una manera sencilla de realizar la estimación. Por estar el cráneo formado por una sustancia rígida, podemos suponer que la longitud de los arcos interiores del mismo se ha mantenido sin cambios. Por otra parte, la morfología general del cerebro de *Proconsul* es similar a la de los monos catarrinos (llamados también monos del Viejo Mundo), y nosotros disponíamos de una serie de moldes internos de cráneos pertenecientes a este último grupo de animales. Las mediciones de los arcos craneales realizadas sobre los moldes mostraron que, en los monos del Viejo Mundo, existe una relación estadística entre la longitud del arco y la capacidad craneal. En el cráneo de *Proconsul* medimos la longitud del arco que se extiende entre el frente del cráneo y el extremo posterior del foramen magno. Admitiendo que en *Proconsul* se cumpliría la misma relación observada en los monos catarrinos, concluimos que la capacidad craneal del fósil estaría comprendida entre 154 y 180 centímetros cúbicos, con una mayor probabilidad para 167 centímetros cúbicos.

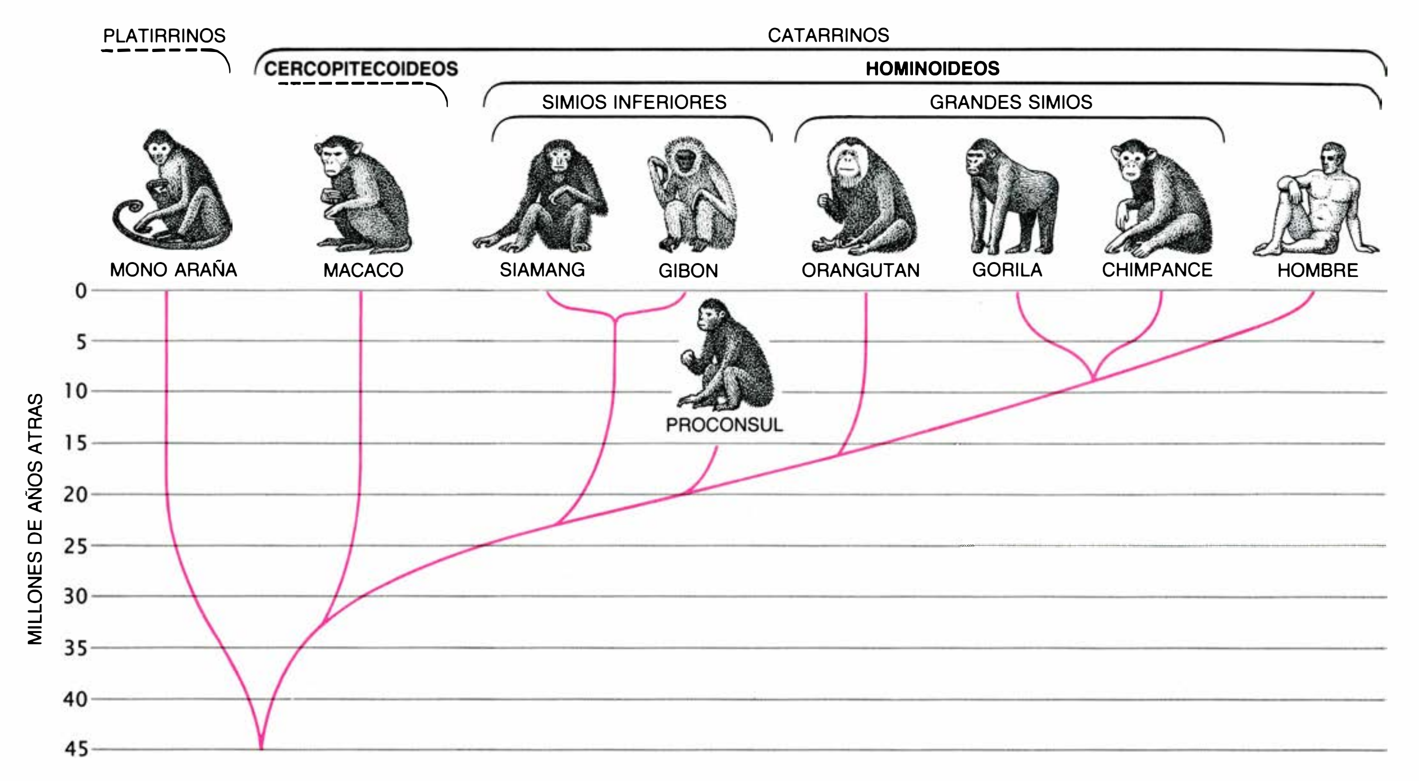
El peso del cuerpo podía calcularse midiendo varios huesos de las extremidades. Realizadas las mediciones oportunas en huesos de este tipo descubiertos en 1951 y en algún otro que se encontraba a nuestra disposición, llegamos a la conclusión de que *Proconsul africanus* estaba más encefalizado que los monos modernos de talla comparable a la suya. Pensamos que esa notable encefalización constituye un rasgo característico de los grandes simios actuales, si bien no podemos afirmarlo con seguridad. Los grandes simios modernos, entre los que se incluye el orangután, el chimpancé y el gorila, son mucho mayores que las especies extintas, y aunque el tamaño de su cerebro resulta en términos absolutos más elevado que el de aquéllas, comparado con el peso del cuerpo es menor en términos relativos. No obstante, si la encefalización define a los grandes simios, este rasgo debió desarrollarse en épocas precoces de su historia evolutiva.

El cráneo de 1948 acaba de revelar otra conexión más entre los primates modernos y los primitivos. Sir Wilfrid Le Gros Clark, a quien debemos la descripción original del cráneo, observó que éste presentaba un seno en la parte interna del hueso frontal. Se trata de una característica importante: los senos frontales aparecen en los seres humanos y en los simios africanos, pero no existen en los orangutanes ni tampoco en los monos del Viejo Mundo o en los simios inferiores (siamangs y gibones). Hace unos tres años, Steven C. Ward, de la Universidad estatal de Kent, realizó un estudio de la forma y dimensiones de los senos faciales en los primates





2. EN LA ISLA RUSINGA se han producido gran parte de los hallazgos de *Proconsul*. Mary Leakey encontró el cráneo de 1948 en el yacimiento R106. El ejemplar de 1951 fue descubierto en el punto R114 por Tom Whitworth. El yacimiento de Kaswanga, descubierto durante la primera campaña de los autores del artículo, ha proporcionado cientos de fósiles de primates, incluyendo al menos nueve esqueletos de *Proconsul*.



3. ARBOL EVOLUTIVO de los hominoideos. Los hominoideos son la superfamilia de primates que incluye los simios inferiores (siamangs y gibones), los grandes simios (chimpancés, gorilas y orangutanes) y el hombre. *Proconsul* vivió hace 18 millones de años, durante el Mioceno Inferior. Aunque comparte algunos caracteres con los gorilas y chimpancés modernos, la mayoría de estos rasgos son bastante generales, sin especializar; ello sugiere que *Proconsul* representa el último antepasado común a grandes simios y hombre. (La "recomposición" de *Proconsul* pone de relieve viejos defectos de la paleontología.)

superiores. Durante ese estudio examinó el cráneo de *Proconsul*. Desgraciadamente, se había rellenado con escayola la región del seno y sólo podía observarse con rayos X. Ahora bien, a causa del peculiar modo de fosilización de los organismos incluidos en los sedimentos de la isla Rusinga, no fue posible, por este procedimiento, distinguir la matriz rocosa que rellenaba el seno de lo que era el hueso en sí y de la escayola. Ward se vio así incapaz de verificar la observación de Le Gros.

Hasta que uno de nosotros (Walker) logró desprender la escayola. Dejábase patente un seno frontal, relleno de matriz y particularmente grande, que se prolongaba hacia la parte posterior, confirmando de ese modo que estos simios del Mioceno tenían, al menos, algunas afinidades con los grandes simios actuales y que, en ese aspecto, se alejaban de los simios inferiores y de los monos del Viejo Mundo.

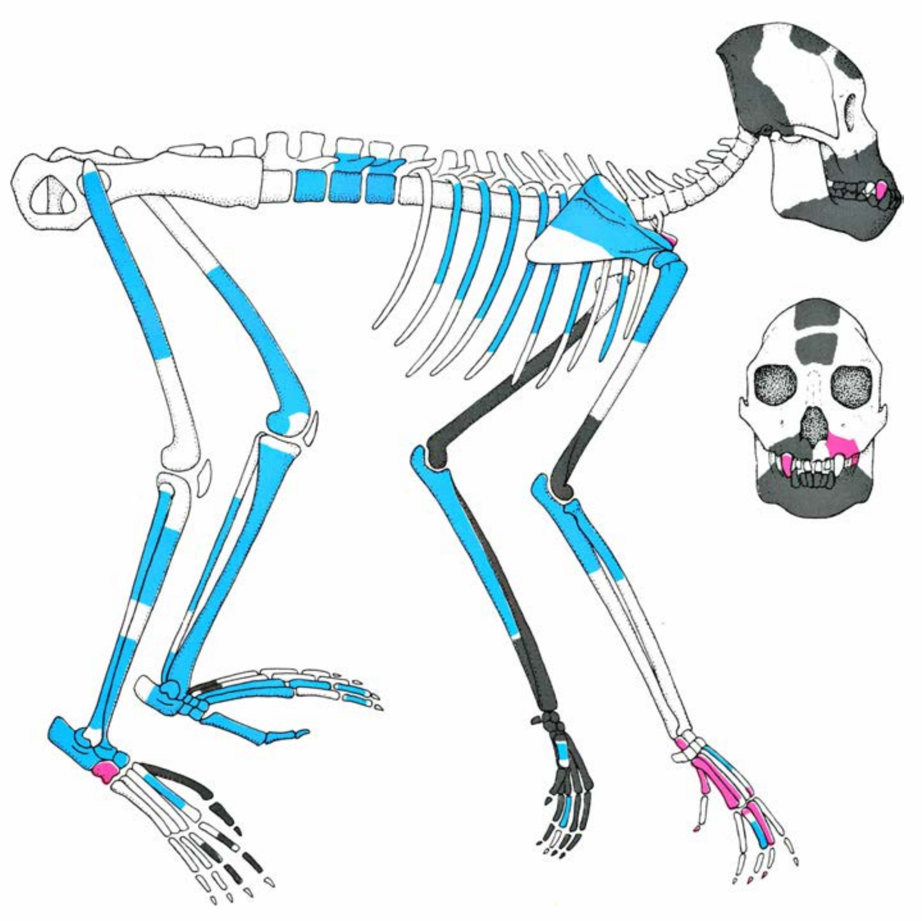
Más complicada aún resulta la historia del especimen de *Proconsul* hallado en 1951. Cuando se descubrió este ejemplar en la localidad denominada R114, los huesos, nada más ser extraídos de la dura roca verde en la que se encontraban incluidos, fueron enviados a Le Gros Clark, quien a su vez los pasó a John R. Napier. Napier y su colega Peter R. Davis, ambos del Royal Free Hospital de Londres, utilizaron los huesos para esbozar un esquema de la evolución de la mano y la extremidad del primate y escribieron una monografía sobre el tema que se convirtió en una pieza clásica. En 1964 Napier devolvió los huesos a Kenya.

Los huesos volvieron a reunirse con el resto de *Proconsul* gracias a un suceso fortuito. En los primeros años de la década de 1980, una colección de restos de fósiles de cerdos, dejados en préstamo a un paleontólogo, fue devuelta al Museo Nacional de Kenya. La colección incluía un esqueleto fósil de cerdo procedente de la sima, así como un bloque de roca verde que contenía varios huesos articulados, atribuidos también a un cerdo. Sin embargo, el especialista en suidos había observado que estos últimos no eran huesos de especies porcinas y los había separado del resto de la colección.

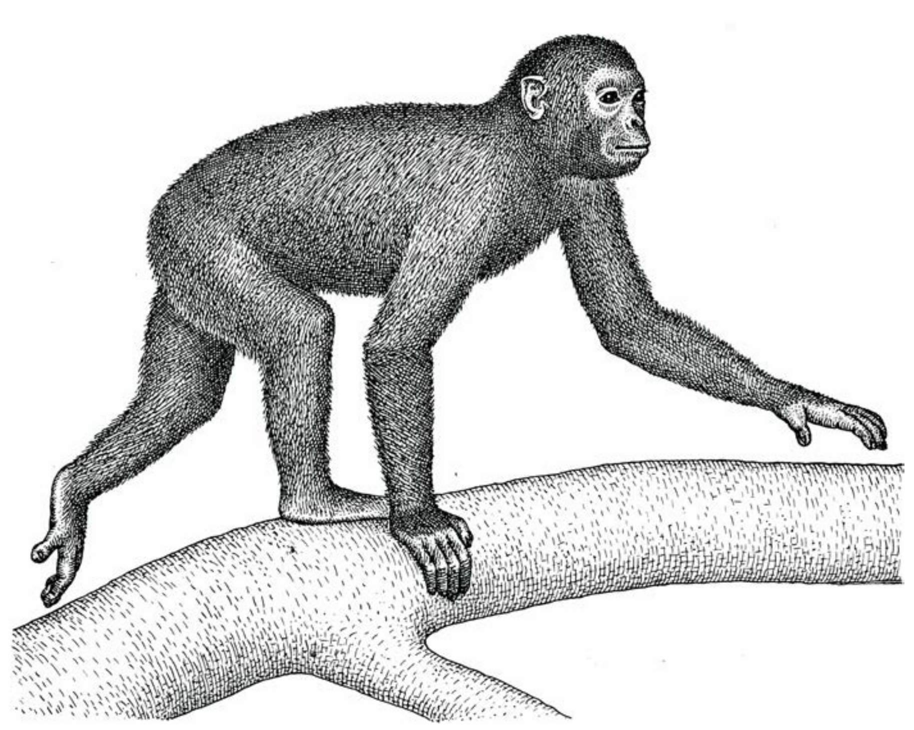
Durante una visita al Museo, en 1980, uno de nosotros (Walker) examinó los fósiles y se dio cuenta de que eran huesos de la pierna y el pie del mismo individuo de *Proconsul* descrito por Napier y Davis. La identificación resultó más fácil gracias a que se trataba de un individuo subadulto: muchas de las epífisis —los extremos de un hueso en expansión, separados de sus diáfisis por cartílagos de crecimiento— tenían todavía que fusionarse con los huesos propiamente dichos; tal era exactamente el estado en que se encontraba el especimen de Napier. Los nuevos huesos permitían también a los investigadores ver, por primera vez, cuáles eran las proporciones de la mano y el pie en *Proconsul*; éstos poseen, a un tiempo, rasgos de simio y mono y muestran que *Proconsul* fue un cuadrúpedo de movimientos lentos.

La devolución de los restos fósiles de cerdo llevó a Pickford y a uno de nosotros (Walker) a sospechar que podrían haber otros huesos de *Proconsul* extraviados por los cajones del Museo de Kenya. Más. Cuando ya se habían recuperado todos los huesos de la localidad excavada, siguieron saliendo a la luz nuevas piezas de *Proconsul*: una escápula, fragmentos de un húmero, más huesos de la mano, los cúbitos, las tibias y trozos de ambos fémures. Con todo esto ya era posible realizar una estimación de las propiedades de las extremidades del animal, las cuales revelan, de nuevo, tanto rasgos de simio (el miembro anterior) como de mono (la pierna).

Si los huesos de un simio y los de un cerdo habíanse mezclado en el cajón de un museo, lo mismo pudo haber ocurrido durante los primeros momentos de la excavación. Pickford y uno de nosotros (Walker) se dispusieron a investigar los huesos de *Proconsul* que quizá habían pasado inadvertidos en la localidad R114, sin tocar desde 1951 y cubierta ya por la vegetación. Ambos in-



4. COMPONER PIEZA A PIEZA el esqueleto de *Proconsul* ha costado 30 años. Las partes del descubrimiento inicial, realizado por Whitworth en 1951, están sombreadas en gris. Los fragmentos teñidos de azul fueron encontradas por Pickford y Walker en 1984, entre las colecciones del Museo Nacional de Kenya. Las piezas pintadas en rojo corresponden a los hallazgos realizados por los autores, en 1984, en el yacimiento R114.



5. UNA RECONSTRUCCION del aspecto que podría haber tenido *Proconsul* hace 18 millones de años. La menor de las dos especies que vivieron en la isla Rusinga tenía la talla de una hembra de babuino, con los miembros anteriores y posteriores de aproximadamente la misma longitud. Era un animal de movimientos lentos que no había desarrollado ninguna adaptación para el salto, braquiación, desplazamiento sobre los nudillos o vida sobre el suelo. La especie grande pesaba casi cuatro veces más que la especie pequeña.

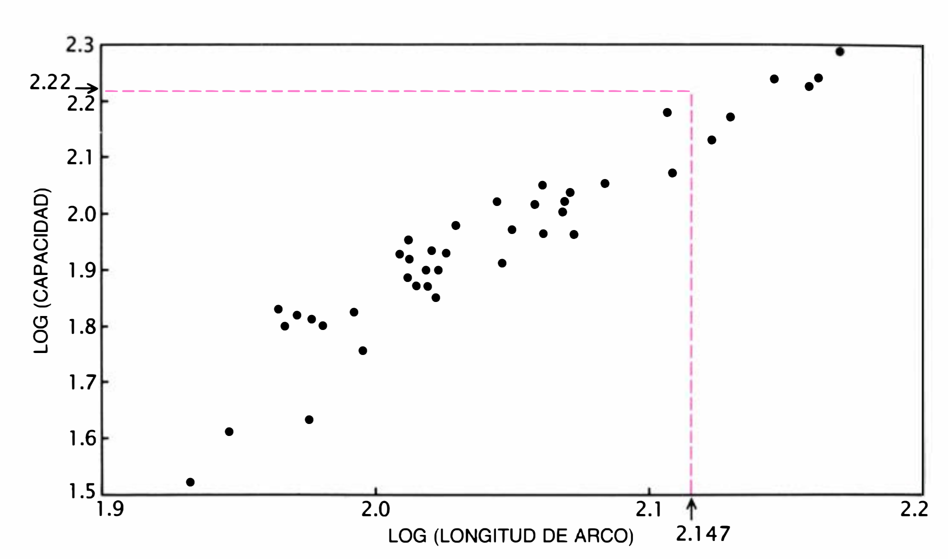
vestigadores hallaron el lugar, pero no tuvieron tiempo para excavar. La excavación realizada por nosotros en 1948 en la localidad R114 fue recompensada por el hallazgo del maxilar que faltaba, media clavícula, un diente canino, un importante hueso del pie y un bloque que contenía huesos de la mano derecha, sin contar diversos restos más de otros mamíferos. De este modo, después de más de treinta años, se volvieron a juntar los numerosos y variados fragmentos de *Proconsul africanus*, rescatados tanto de la sima en la que se encontraban todos inicialmente como de los arcones del museo adonde algunos de ellos habían ido a parar equivocadamente etiquetados. El esqueleto de "1951" es ahora el más completo entre todos los pertenecientes a hominoideos de gran talla del Mioceno.

La campaña de 1984 en R114 suministró también importante información acerca de la estructura de la sima en sí, información que a su vez incidiría en un mejor conocimiento de la paleoecología de *Proconsul*. Varias semanas de excavación en la roca sedimentaria vecina a la sima mostraron que ésta tenía, al menos, cuatro metros de profundidad. Las capas de limos no se extienden simétricamente alrededor de ella, sino que, por lo que se ve, se depositaron de manera desigual alrededor de un objeto que se encontraba en posición vertical allí donde existe hoy la chimenea de la roca verde. Las pequeñas fallas que hay a cada lado de la sima confirman esta suposición: algo se había interpuesto en el camino de los limos que eran arrastrados por el agua, cuando éstos se depositaban, y con el tiempo ese algo colapsó provocando la fracturación de los sedimentos. Esta hipótesis explicaba también la disposición asimétrica de tales sedimentos, en la zona situada aguas arriba eran más numerosos los cantos y las arenas que en la parte que quedaba aguas abajo.

Con el tiempo se comprobó que la sima en cuestión no era tal. Había ocurrido que un árbol potente había crecido allí hace 18 millones de años y había quedado parcialmente enterrado por limos y arenas. El árbol aguantó el enterramiento; tras su muerte, por esa causa probablemente, no tardaría en ahuecarse su interior. Una vez vaciado, comenzó a ser ocupado por lagartos, pitones, murciélagos y pequeños carnívoros que dejaron sus huesos en el tronco, y también los de sus presas. Probablemente, uno de los carnívoros atrapó a *Proconsul* y lo tansportó hasta el interior del árbol para devorarlo allí;

de hecho, ciertas articulaciones de este ejemplar de *Proconsul* muestran señales de dientes. Con el tiempo, el árbol se rellenaría con una mezcla de huesos y sedimentos que más adelante solidificaría hasta formar la chimenea de roca que ha llegado a la actualidad.

La solución del misterio de la sima resolvió dos importantes problemas. En primer lugar, sabemos ya con certeza que los fósiles y el relleno verde del depósito son de la misma edad, geológicamente hablando, que los sedimentos circundantes de limos y arenas; estos últimos se han fechado aplicando el método del potasio-argón en las lavas que se encuentran por encima y por debajo de ellos, obteniéndose una edad de unos 18 millones de años. En segundo lugar, gracias al árbol, que contiene fósiles de animales que usaron su tronco de guarida o refugio, esta asociación faunística representa la comunidad local de *Proconsul*.



6. CAPACIDAD CRANEAL de los actuales monos del Viejo Mundo y su relación con la longitud del arco del cráneo. En el cráneo de *Proconsul* de 1948 fue imposible realizar mediciones directas de la capacidad cerebral debido a que ese cráneo ha sufrido aplastamiento (*véase la figura 1*). Sin embargo, la longitud del arco del cráneo no experimenta cambios con la deformación. Además, el cerebro de *Proconsul* debió ser similar, en su forma, al cerebro de los monos del Viejo Mundo, lo que permite dar por supuesto que aquél mantuviera la misma relación entre longitud del arco y capacidad craneal. El cráneo de 1948 presentaba una longitud del arco interno, desde la frente al foramen magno (el gran orificio situado en la base del cráneo), de unos 140 milímetros. Representando el logaritmo de ese valor sobre el eje horizontal se puede deducir del eje vertical una capacidad craneal para *Proconsul* cifrada en apoximadamente 167 centímetros cúbicos.

La excavación de la localidad R114 condujo por accidente a otro hallazgo importante. Una de las frecuentes tormentas eléctricas que se suceden sobre el lago Victoria inundó la excavación R114. Imposibilitados para trabajar en la sima, el equipo del Museo Nacional de Kenya, dirigido por Kamoya Kimeu, aprovechó para trabajar en las áreas vecinas. La campaña alternativa resultó un éxito: se descubrió una localidad que ha proporcionado desde entonces cientos de huesos enteros y miles de fragmentos. Es más, la nueva localidad, rebautizada con el nombre de Yacimiento de Primates de Kaswanga, se ha convertido en un auténtico filón paleontológico, por lo que hemos negociado un acuerdo por cinco años con el cacique de la isla Rusinda para evitar que la zona se destine al cultivo de maíz.

Entre los hallazgos más importantes, contamos, por lo menos, nueve esqueletos, parciales o completos, de *Proconsul* que habían sido arrastrados hasta un pequeño barranco en los terrenos del Mioceno Inferior. Pertenecen a simios cuyas edades abarcan desde las correspondientes a crías muy pequeñas hasta ejemplares adultos; es probable la presencia de machos y hembras. Se conocen ya todos, o casi todos, los elementos esqueléticos de *Proconsul*, procedentes de uno o varios individuos.

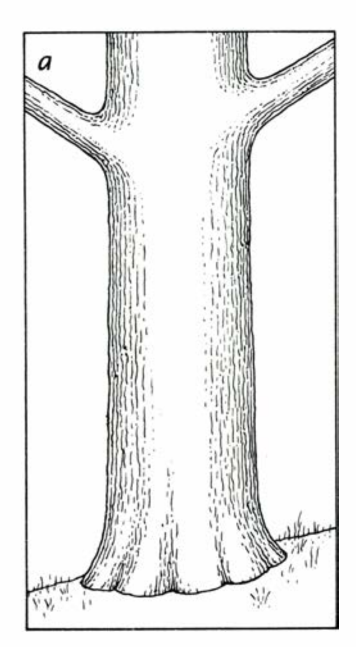
Para nuestra desgracia, la mayoría de los huesos de Kaswanga han quedado rotos o astillados por culpa de la compactación de las arcillas y limos en los que han permanecido sepultados, o por efecto de las raíces de vegetales. Esto nos ha obligado a recomponer los ejemplares con pegamento, proceso que aún continúa en el momento actual. Aunque hemos tenido bastante éxito con los huesos grandes, costó distinguir los pequeños fragmentos cilíndricos de los principales huesos de extremidades de las crías, de los huesos de los dedos de la mano y el pie correspondientes a juveniles y adultos.

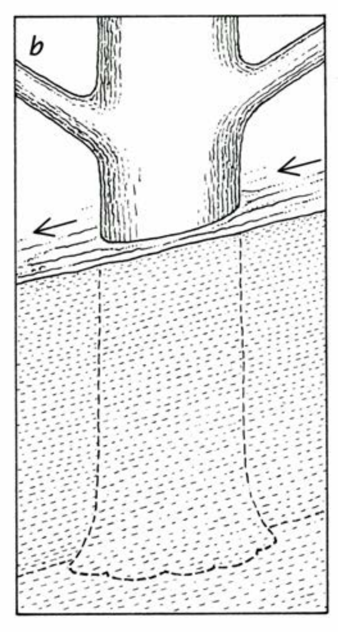
Los fósiles de Kaswanga nos han permitido calcular las proporciones de la mano, el pie y las extremidades de *Proconsul*; cuando terminemos los trabajos de encolado, esperamos establecer los diversos estadios de crecimiento de *Proconsul*. El análisis de los nuevos huesos nos ocupará algunos años; por eso hemos optado por concentrar primero nuestros esfuerzos en los huesos que han sido objeto de investigación provisional. Y actuando en esa línea hemos deshecho el nudo de una controversia relativa a la anatomía funcional y el significado evolutivo de la muñeca de *Proconsul*.

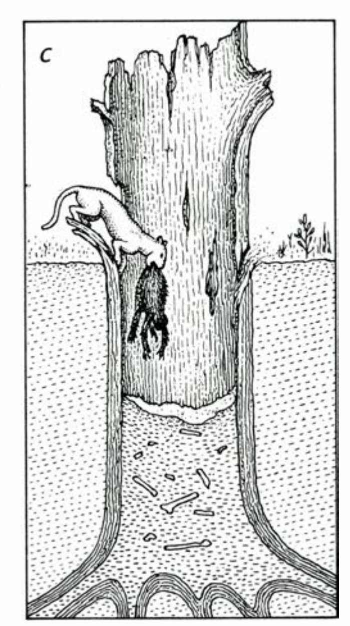
La muñeca de simios y humanos difiere de la propia de la mayoría de los mamíferos, incluidos los monos del Viejo Mundo: no existe articulación estricta entre uno de los huesos del brazo —el cúbito— y la muñeca en sí. En los monos del Viejo Mundo, en cambio, encontramos una pequeña extensión ósea que surge del cúbito y se apoya en una cavidad que dejan los dos huesos de la muñeca, el pisiforme y el piramidal.

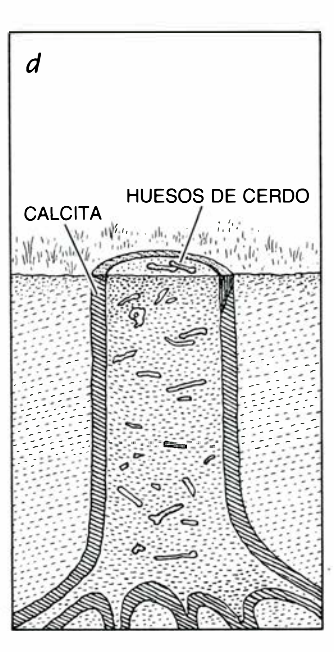
De los huesos del esqueleto de *Proconsul* hallado en 1951, los correspondientes a la mano izquierda han dado origen a múltiples artículos científicos, amén de varias tesis doctorales. Sin embargo, el piramidal ha sufrido graves daños que han sido fuente de discrepancias acerca de la naturaleza de las articulaciones de la muñeca de *Proconsul*. Como resultado de los hallazgos de Kaswanga, disponemos ya de varios ejemplares nuevos del pisiforme y el piramidal. En un trabajo realizado en colaboración con K. Christopher Beard, de la Universidad Johns Hopkins, hemos demostrado que las muñecas de *Proconsul* y de los monos del Viejo Mundo son similares en el sentido de que, en ambos, existe una articulación directa entre muñeca y cúbito. Aunque en este aspecto la muñeca de *Proconsul* puede considerarse primitiva, en otros prefigura la mayor movilidad que presentan ciertas partes de las muñecas de hominoideos actuales.

Este modelo "híbrido", o en mosaico, suele hacerse manifiesto cuando se analizan los hominoideos del Mioceno. Tal como ha señalado Michael D. Rose, de las facultades de medicina y odontología de Nueva Jersey en Newark, los hominoideos del Mioceno ni son semejantes a los monos del Viejo Mundo, ni tampoco a los grandes simios; sólo se parecen a los hominoideos del Mioceno. Esta combinación única de caracteres convierte la interpretación funcional de los fósiles de *Proconsul* en una tarea ardua: además









7. HOYO donde se desenterró el ejemplar de *Proconsul* de 1951; corresponde a los restos fosilizados del tronco de un árbol. Durante el Mioceno Inferior un árbol que se levantaba cerca de un río (*a*) fue enterrado gradualmente por sedimentos grises (*b*). El árbol se pudrió y quedó hueco (*c*), y así los animales pudieron usarlo de guarida. Probablemente un creodonto transportó un cadáver de *Proconsul* hasta el interior del tronco. Andando el tiempo, el tronco se rellenó con huesos y con los sedimentos verdes en los que quedaron englobados los fósiles hasta 1951 (*d*). La corteza del árbol fue reemplazada por calcita.

de no existir ningún animal moderno que sirva de modelo para explicar la anatomía de *Proconsul*, cada complejo anatómico del mismo muestra una combinación única de caracteres. Así, algunos huesos del tobillo, estilizados, semejan los de los monos; en cambio, el pulgar del pie es robusto y comparable con el de los primates. El mismo modelo híbrido se hace patente en la pelvis de *Proconsul*: el ilion, o parte superior de la misma, es similar al de los monos del Viejo Mundo, mientras que el acetábulo (el punto de articulación de la cabeza del fémur) es grande y plano, como en los antropoides.

Siguiendo varias líneas de trabajo, los paleontólogos empiezan a sospechar que los antepasados de los monos del Viejo Mundo debieron haber pasado más tiempo sobre el suelo que sus actuales descendientes. Esta hipótesis permite entender que muchos de los rasgos que *Proconsul* comparte con los grandes simios actuales sean rasgos primitivos, que estos animales han retenido durante millones de años. Esas características sin función especializada tampoco son signo de ningún parentesco evolutivo. (Después de todo, tanto los seres humanos como las zarigüeyas tienen cinco dedos.) La pelvis de *Proconsul*, por ejemplo, muestra que estas criaturas no poseían callosidades isquiátricas o rugosidades de las nalgas sobre las que se sientan los gibones y los monos del Viejo Mundo. La ausencia de callosidades, sin embargo, es una condición primitiva; la presencia de estas estructuras se considera una especialización que sirve para unir gibones y monos del Viejo Mundo, si bien la falta de las mismas no enlaza necesariamente a *Proconsul* con los chimpancés, que carecen también de callosidades isquiátricas. Simplemente, el chimpancé ha retenido una condición ancestral, que procede de tiempos anteriores a la época de *Proconsul*. Por otra parte, el seno frontal está presente en *Proconsul*, seres humanos y grandes simios, pero no en gibones ni en los monos. Estas especializaciones compartidas indican que *Proconsul* se relaciona más estrechamente con los grandes simios actuales que con los monos modernos.

Nuestras recientes excavaciones han arrojado también alguna luz sobre las diversas especies de *Proconsul* que vivieron en las islas Rusinga y Mfangano. Durante algún tiempo, varios investigadores sostuvieron que la especie pequeña de *Proconsul*, hallada en las islas y representada por el cráneo encontrado por Mary Leakey en 1948, no difería de *Proconsul africanus*, la forma descubierta en principio en tierra firme y representada por las mandíbulas desenterradas por Hopwood. Si este punto de vista fuera correcto, deberían haber existido al menos tres especies, dos pequeñas y otra grande, *Proconsul nyanzae*, que se encuentra también en las islas. Otros investigadores, como Pickford, Jay Kelley, de la Universidad de Brown, y David Pilbeam, de la Universidad de Harvard, se inclinan por la existencia de sólo dos especies: los especímenes pequeños encontrados en las islas serían hembras pertenecientes a la especie de mayor talla, *Proconsul nyanzae*.

Junto con Christopher B. Ruff, adscrito a nuestro departamento de la facultad de medicina de la Universidad Johns Hopkins, hemos sometido a prueba esa última hipótesis midiendo la talla corporal de los especímenes grandes y pequeños y comparando las diferencias halladas con las registradas de machos y hembras de primates actuales. Entre los cientos de nuevos fósiles encontrados en el yacimiento de Kaswanga hay varios fémures de *Proconsul* de ambas tallas. Ruff ha descubierto que se puede estimar de modo muy preciso el peso del cuerpo de los primates cuadrúpedos a partir de mediciones de la distribución de la sustancia ósea observada en cortes del fémur. Dos fémures grandes y dos pequeños, pertenecientes a *Proconsul* y todos ellos completos (y que en la medida en la que podemos valorarlo, apoyándonos en los especímenes disponibles, son bastante representativos) ha proporcionado una estimación del peso del cuerpo que indica, respectivamente, 37 y 9,6 kilogramos, para una y otra talla. Esto es, las supuestas hembras pesaban la cuarta parte de los supuestos machos.

Ningún mamífero terrestre, sin mencionar los primates, muestra un dimorfismo sexual tan extremo en el peso del

cuerpo. Esto nos condujo a rechazar la idea de que las formas grandes y pequeñas de *Proconsul* halladas en la isla Rusinga representasen machos y hembras de la misma especie. Llegamos, por tanto, a la conclusión de que fueron dos las especies de simios primitivos que habían vivido en la isla Rusinga y en Mfangano hace 18 millones de años: una especie pequeña, que difiere de la especie hallada por Hopwood en tierra firme, y la especie de mayor talla, *Proconsul nyanzae*. Parecen haber sido similares entre sí en forma y proporciones, pero muy diferentes en cuanto a talla corporal.

Las semejanzas de forma y proporciones evidencian una afinidad importante en la postura y el modo de locomoción. Los especímenes insulares de *Proconsul* exhiben un diseño primitivo en las proporciones de las extremidades y en la superficie de las articulaciones; los miembros anteriores y posteriores, por ejemplo, presentaban una longitud similar. Por otro lado, ambas formas carecen de cola, una condición especializada. En su monografía clásica, Napier y Davis, basándose solamente en el análisis de los miembros anteriores, llegaron a la conclusión de que *Proconsul* fue un activo cuadrúpedo saltador, que se movía como los actuales langures asiáticos, amén de poseer rasgos esqueléticos que manifestaban hábitos de braquiación.

Los nuevos fósiles revelan, por el contrario, que las especies de *Proconsul* fueron arborícolas, de movimientos bastante lentos, probablemente cautelosos y que no presentaban especializaciones obvias para el salto, la braquiación, el desplazamiento sobre los nudillos o la vida en el suelo.

El *status* de *Proconsul* ha cambiado considerablemente en los 60 años que han transcurrido desde que Gordon encontró el primer fragmento de mandíbula. Hopwood pensaba que *Proconsul* era un antepasado del chimpancé, idea que se afianzó, más allá de los años sesenta y setenta, entre algunos paleontólogos que consideraban las diferentes formas de *Proconsul* como ancestrales de las distintas especies de grandes simios actuales.

Pero la última década ha sido fecunda en descubrimientos de nuevo material, procedente no sólo de la isla Rusinga. Así, Richard y Maeve Leakey, del Museo Nacional de Kenya, han encontrado por lo menos tres nuevos géneros de simios, con una antigüedad de 17 millones de años, localizados en un yacimiento situado al norte de Kenya. Los nuevos fósiles presentan algunas semejanzas con *Proconsul* y varias diferencias. Sin lugar a dudas, la diversidad de simios que hubo, durante el Mioceno Inferior, en aquella pequeña parte del Africa Oriental, fue mucho mayor de lo que en principio se había pensado.

Estos descubrimientos han servido para poner de relieve que las interpretaciones tradicionales sobre la evolución de los hominoideos primitivos pecaron de una grosera hipersimplificación, que se fundaba en muestras limitadas en espacio y tiempo. La comprobación de ese hecho, unido a la observación de que muchos caracteres atribuidos a los hominoideos podían ser, en realidad, caracteres primitivos del grupo de los antropoides, ha proporcionado a los investigadores una imagen muy distinta de *Proconsul*.

Proconsul no fue un antecesor especializado del moderno chimpancé o del gorila. La verdad es que presenta pocos rasgos especiales que le relacionen con estos primates actuales. *Proconsul* aparece como un antepasado general de todos los grandes simios y de los seres humanos. Tan general que, si lo comparamos con un antepasado teórico de todos los hominoideos, anterior a él, solamente algunos rasgos, muy pocos —como el seno frontal o la ausencia de callosidades isquiátricas— revelarían que este primate del Mioceno Inferior vivió con posterioridad a haberse producido la divergencia entre los simios inferiores y los superiores.

BIBLIOGRAFIA COMPLEMENTARIA

THE FORE-LIMB SKELETON AND ASSOCIATED REMAINS OF *PROCONSUL AFRICANUS*. John R. Napier y Peter R. Davis en *Fossil Mammals of Africa*, vol. 16, págs. 1-69; 1959.

MIOCENE HOMINOID POSTCRANIAL MORPHOLOGY. Michael D. Rose en *New Interpretations of Ape and Human Ancestry*, editado por Russell L. Ciochon y Robert S. Corruccini. Plenum Press, 1983.

NEW POSTCRANIAL FOSSILS OF *PROCONSUL AFRICANUS* AND *PROCONSUL NYANZAE*. Alan Walker y Martin Pickford en *New Interpretations of Ape and Human Ancestry*, editado por Russell L. Ciochon y Robert S. Corruccini. Plenum Press, 1983.

THE SKULL OF *PROCONSUL AFRICANUS*: RECONSTRUCTION AND CRANIAL CAPACITY. Alan Walker, Dean Falk, Richard Smith y Martin Pickford en *Nature*, vol. 305, n.º 5934, págs. 525-527; 6 de octubre de 1983.

NEW WRIST BONES OF *PROCONSUL AFRICANUS* AND *P. NYANZAE* FROM RUSINGA ISLAND, KENYA. K. Christopher Beard, Mark F. Teaford y Alan Walker en *Folia Primatologica*, vol. 47, n.º 2, págs. 97-118; 1986.

El problema de la red mínima

¿Cuál es la red lineal de longitud total mínima que interconecta un conjunto arbitrario de 100 puntos, pongamos por caso? La solución de este problema se ha resistido a los ordenadores más rápidos y a los matemáticos más sagaces

Marshall W. Bern y Ronald L. Graham

Una compañía imaginaria, la Cía. Telefónica Steiner, estimó que podría ahorrar cientos de millones de pesetas si lograse determinar la red mínima de tendido telefónico capaz de interconectar a sus 100 clientes. Tratando de dar con una solución, la Cía. Steiner contrató los servicios de Informática Cavalieri, de nombradía mundial por la eficacia de sus programadores y la rapidez de sus equipos de cómputo. Una semana después, la Cavalieri presentaba a la Cía. Steiner un programa para resolver su problema, demostrando que el programa había hallado en sólo una hora la red mínima correspondiente a 15 de los abonados de ésta. La Cía. Steiner abonó a la Cavalieri 100.000 pesetas por el programa y se comprometió a pagar una peseta por cada segundo de ordenador invertido en la determinación de la solución completa. Pues bien, para cuando el ordenador hubiera concluido el cómputo correspondiente a los 100 abonados, la compañía telefónica adeudaría cientos de billones de pesetas en concepto de tiempo de cómputo y sus abonados se habrían dispersado, ¡fuera por voluntad propia o a causa de la deriva de los continentes!

¿Acaso era defectuoso el programa que Cavalieri preparó para la Cía. Steiner? El dilema anterior es uno de los ejemplos del problema de Steiner, que se plantea hallar la red de segmentos rectilíneos de mínima longitud total que permita conectar entre sí cierto conjunto de puntos. El problema de Steiner no puede ser resuelto trazando sin más segmentos entre los puntos dados, pero sí añadiendo puntos nuevos, llamados puntos de Steiner, que sirven de nudos de enlace en una red de longitud mínima. Los matemáticos y los especialistas en ciencias de cómputo han ideado algoritmos (esto es, procedimientos estrictamente formulados) con el fin de determinar el número y la ubicación de los puntos de Steiner. Empero, ni aun los mejores de tales algoritmos, instalados y funcionando en los ordenadores más rápidos, alcanzan a proporcionar una solución para conjuntos grandes de puntos arbitrarios, porque el tiempo que habrían de invertir para lograrlo es tan largo que es irrealizable. Además, el problema de Steiner pertenece a una categoría de problemas tales que son muchos los científicos informáticos convencidos de que jamás se podrá hallar para ellos un algoritmo eficiente. Por esta razón, la Informática Cavalieri debería quedar exonerada de culpa.

Por otra parte, la Cavalieri sí podría haber preparado un programa práctico que proporcionase soluciones aproximadas, si bien algo más largas que una red mínima. Son hoy muchas las aplicaciones del problema de la red mínima para las que se están calculando de forma rutinaria soluciones aproximadas; cabe citar el diseño de circuitos integrados, la determinación del árbol filogenético de un grupo de organismos y la minimización del material utilizado en redes de líneas telefónicas, oleoductos y carreteras.

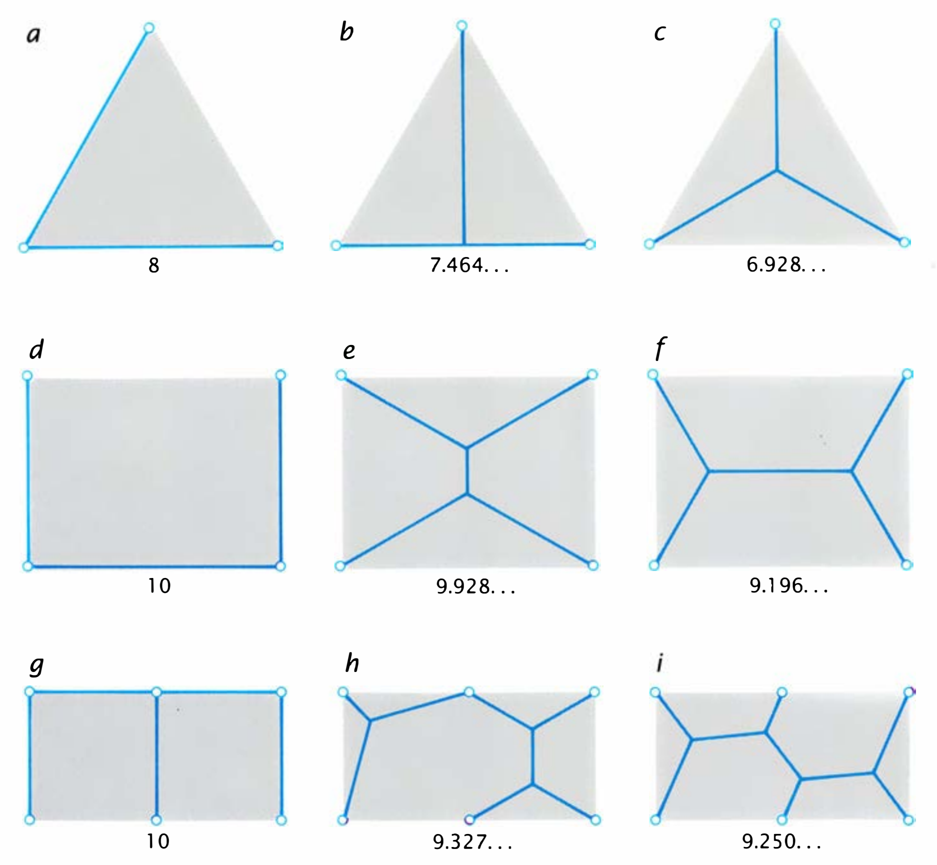
En su forma general, el problema de Steiner apareció publicado en 1934, en un artículo de Miloš Kössler y Vojtěch Jarník, pero el problema no alcanzó difusión y popularidad hasta 1941, al darle cabida Richard Courant y Herbert E. Robbins en su libro *What is Mathematics*? Courant y Robbins ligaban el problema a un trabajo de Jakob Steiner, que fue profesor de matemáticas de la Universidad de Berlín a mediados del siglo pasado. El trabajo de Steiner buscaba determinar un único punto cuyos enlaces con un conjunto de puntos dados tuviesen longitud total mínima. Pero hacia 1640 se planteó ya un caso particular de ambos problemas (el problema en que trabajaba Steiner y el que hoy lleva su nombre), a saber: hallar el punto P que hace mínima la suma de las distancias desde P a tres puntos dados. Evangelista Torricelli y Francesco Cavalieri resolvieron independientemente tal problema. Ambos dedujeron que, si los ángulos en P son todos de 120 grados o más, la suma de distancias se minimiza.

Sabiendo que los ángulos en P han de medir al menos 120 grados, Torricelli y Cavalieri idearon una construcción geométrica que permitiera la determinación de P [*véase la figura 3*]. Se trazan segmentos rectilíneos que conecten los puntos dados (llamémoslos A, B y C, con B ocupando el vértice del mayor de los ángulos). Cuando el ángulo B es mayor o igual que 120 grados, el punto P coincide con el punto B. Dicho de otro modo, la red mínima se reduce en este caso a los dos segmentos rectilíneos que unen A con B, y B con C. Cuando el ángulo en B es menor que 120 grados, el punto P ha de encontrarse en el interior del triángulo. Para determinar dónde, se construye un triángulo equilátero sobre el lado mayor del triángulo, o sea, entre los puntos A y C. El tercer vértice del

MARSHALL W. BERN y RONALD L. GRAHAM llevan muchos años estudiando problemas de redes mínimas. Bern trabaja en el Centro de Investigación Palo Alto, de Xerox. Acabado el primer ciclo superior de su formación en la Universidad de Texas en Austin pasó a la división de procesado de señal de la empresa Ensco. Se recibió de doctor en informática por la Universidad de California en Berkeley, en 1987. Graham es director adjunto de investigación de la división de ciencias de información en los Laboratorios AT&T Bell, además de profesor en la Universidad de Rutgers. Se doctoró en matemáticas por Berkeley en 1962.



1. UNA COMPUTADORA DE LAMINAS JABONOSAS (*arriba*) desafía a un ordenador electrónico (*abajo*) en la tarea de hallar la red más corta de segmentos rectilíneos que interconecta 29 ciudades de los Estados Unidos y de Canadá. La computadora de burbujas, en la que los puntos geográficos se remedan mediante clavijas, hace mínima la longitud de las películas jabonosas en un área local. Proporciona un red de intercomunicación corta, aunque no necesariamente la más corta de todas. El ordenador electrónico pone en práctica un algoritmo, ideado por Ernest J. Cockayne y Denton E. Hewgill, adscritos a la Universidad de Victoria, que garantiza la auténtica red mínima. Este problema, que tiene 29 puntos, se encuentra casi en el límite de la capacidad de cómputo actual.



2. EL PROBLEMA DE LA RED MINIMA tiene diversas soluciones en los casos de un triángulo equilátero, un rectángulo y una "escalera". En *a*, *d* y *g*, los puntos se conectan sin adición de otros nuevos, y forman una solución llamada de tendido mínimo o árbol. Vemos árboles de Steiner en *c*, *e*, *f*, *h* e *i*, formados por adición de nodos de enlace. Tan sólo *c*, *f*, e *i* son árboles de Steiner mínimos o redes minimales. El número arábigo que figura bajo cada solución da la longitud total aproximada de los segmentos rectilíneos.

triángulo equilátero, rotulado *X*, se encuentra frente al punto *B*. Se trazan la circunferencia circunscrita al triángulo equilátero y el segmento rectilíneo que une *B* con *X*. El punto *P* se encuentra en la intersección de este segmento y la circunferencia. Al unir los puntos *A*, *B* y *C* se crean tres ángulos de 120 grados exactamente, que proporcionan la red mínima. Además, la longitud del segmento que une *B* con *X* resulta ser igual a la longitud de dicha red. Para los fines del presente artículo, llamaremos a *X* punto de reemplazo, pues al sustituir por *X* los puntos *A* y *C* la longitud de la red no varía.

Son muchas las propiedades que comparten el problema de los tres puntos y el problema pluripuntual de Steiner. La forma de las soluciones, conocidas por árboles, es tal que, al eliminar uno cualquiera de los segmentos rectilíneos de la red mínima, uno de los puntos dados queda desconectado. Con otras palabras, no es posible salir de un punto de la red y retornar a él sin recorrer dos veces un mismo tramo. Por esta razón, tanto las soluciones del problema tripuntual como del pluripuntual se denominan árboles de Steiner; los segmentos se llaman lados y, puntos de Steiner, aquellos puntos (análogos al *P*) que es preciso añadir para construir el árbol.

El problema tripuntual de Steiner proporciona, asimismo, información sobre la geometría de los árboles de Steiner mínimos. En primer lugar, cada ángulo ha de medir al menos 120 grados, lo que implica que ningún punto dado puede estar conectado al árbol por más de tres lados. Segundo, en cada punto de Steiner concurren exactamente tres lados, que forman ángulos de 120 grados. En tercer lugar, el número de lados de un árbol es siempre una unidad menor que el número de puntos dados más el número de puntos de Steiner. Y por último, puesto que en cada punto de Steiner concurren exactamente tres lados y que al menos un lado ha de tocar a cada uno de los puntos dados, el número máximo de puntos de Steiner de un problema arbitrario es dos unidades menor que el número de puntos dados.

Son muchos los distintos árboles de Steiner que se pueden construir para un mismo número y disposición de puntos dados. Algunos de tales árboles, conocidos por soluciones localmente minimales, no pueden ser acortados por una pequeña perturbación como, por ejemplo, un ligero desplazamiento de un lado o la escisión en dos de uno de los puntos de Steiner. Ahora bien, no todo árbol de Steiner localmente minimal determina una de las soluciones mínimas posibles. Pueden resultar necesarias reorganizaciones a gran escala de los puntos de Steiner para transformar una red en un árbol de longitud mínima, en uno de los árboles denominados árboles de Steiner globalmente minimales.

Pongamos de manifiesto lo dicho con un conjunto de puntos dados que definen los cuatro vértices de un rectángulo de cuatro por tres metros. Las soluciones tienen dos puntos de Steiner, que podemos disponer de dos formas distintas. Cada disposición forma un árbol de Steiner, con tres lados conectados a cada uno de los puntos de Steiner en ángulos de 120 grados. Cuando los puntos de Steiner se disponen paralelamente a la altura, el árbol de Steiner localmente minimal que se obtiene posee una longitud total de más de 9,9 metros. Situando los puntos de Steiner paralelamente al lado mayor, se obtiene otro árbol de Steiner localmente minimal, que mide un poco menos de 9,2.

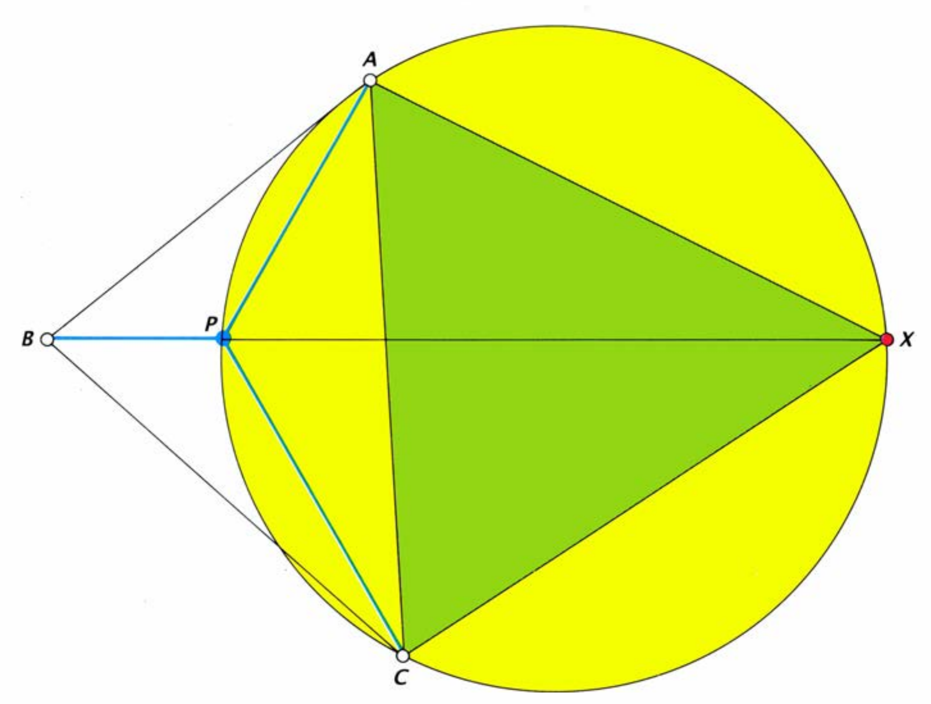
Un método de fuerza bruta para descubrir una red mínima consistiría en determinar, de forma exhaustiva, la totalidad de árboles de Steiner localmente minimales, calcular sus longitudes y elegir el más corto de todos. Ahora bien, como los puntos de Steiner pueden, en principio, estar ubicados en cualquier lugar, no resulta obvio que pueda calcularse en un tiempo finito la totalidad de posibles árboles de Steiner localmente minimales. Z. A. Melzak, de la Universidad de Columbia Británica, superó tal dificultad y puso a punto el primer algoritmo para el problema de Steiner.

El algoritmo de Melzak toma en consideración muchas posibles conexiones entre puntos dados y muchas posibles ubicaciones para los puntos de Steiner. Podemos esbozar tal algoritmo en dos partes. La primera parte se limita sencillamente a distinguir en el conjunto de puntos dados todos los posibles subconjuntos. La segunda crea, para cada subconjunto, cierto número de posibles árboles de Steiner mediante una construcción similar a la utilizada para resolver el problema de los tres puntos.

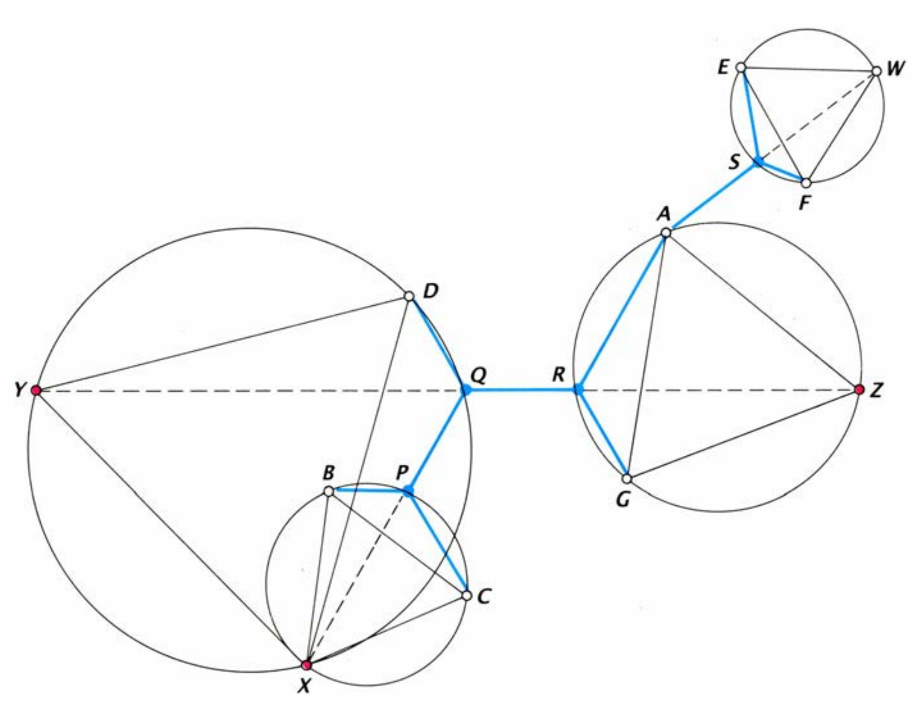
Lo mismo que en el problema de los tres puntos, podemos sustituir dos de los puntos dados por un punto de reemplazo sin modificar la longitud de la solución. Sin embargo, en el problema general, el algoritmo debe tantear qué par ha de ser el reemplazado y, en definitiva, se ve obligado a ensayar todos los pares posibles. Además, el punto de reemplazo puede hallarse situado a uno u otro de los lados del segmento rectilíneo que une el par en cuestión, porque el triángulo equilátero utilizado en la construcción puede orientarse en dos sentidos. Después de reemplazar uno de los pares de puntos del conjunto dado por uno de los dos posibles puntos de reemplazo, cada uno de los pasos ulteriores del algoritmo consiste en lo siguiente: reemplazar dos de los puntos dados, o un punto dado y un punto de reemplazo, o dos puntos de reemplazo, por otro punto de reemplazo, hasta dejar el conjunto reducido a tres puntos.

Una vez determinado el punto de Steiner correspondiente a estos tres puntos, el algoritmo opera retrógradamente, tratando de determinar el punto de Steiner correspondiente a cada punto de reemplazo [*véase la figura 4*]. Una tentativa puede fallar debido a la existencia de restricciones contradictorias que pesan en la ubicación de los puntos de Steiner. Sin embargo, los ensayos que tengan éxito crean un árbol de Steiner que conecta cada punto dado del subconjunto con un lado. Tras examinar todas las secuencias de reemplazo, el algoritmo selecciona para el subconjunto el más corto de tales árboles de Steiner. Combinando de todos los modos posibles los árboles de Steiner mínimos que correspondan a subconjuntos y que interconecten el conjunto de puntos inicialmente dado, se obtendrá la totalidad de árboles de Steiner localmente minimales posible; llegados aquí podremos entonces determinar la geometría de la red mínima global.

El algoritmo de Melzak puede exigir tiempos enormes, incluso en problemas pequeños, por ser muy grande el número de posibilidades que toma en consideración. Un problema de 10 puntos, pongamos por caso, puede descomponerse en 512 subconjuntos de dos o más puntos dados. Aunque los subconjuntos de dos puntos apenas si requieren trabajo, cada uno de los 45 subconjuntos de ocho puntos tiene dos millones de secuencias de reemplazo. Por otra parte, hay más de 18.000 ma-



3. RED MINIMA correspondiente a tres puntos *A*, *B* y *C*; no es difícil de construir. Se levanta el triángulo equilátero *ACX* (*verde*) sobre el lado mayor del triángulo *ABC* y seguidamente se le circunscribe un círculo (*amarillo*). La intersección del círculo y de un segmento rectilíneo trazado desde *B* a *X*, el tercer vértice del triángulo equilátero, define el punto *P*, conocido como punto de Steiner. Al unir con *P* los puntos *A*, *B*, y *C* se forman tres ángulos de 120 grados, por lo que *P* proporciona la red mínima. La longitud del segmento rectilíneo *BX* es igual a la longitud de la red. (Los dibujos del artículo han sido realizados por Gabor Kiss.)



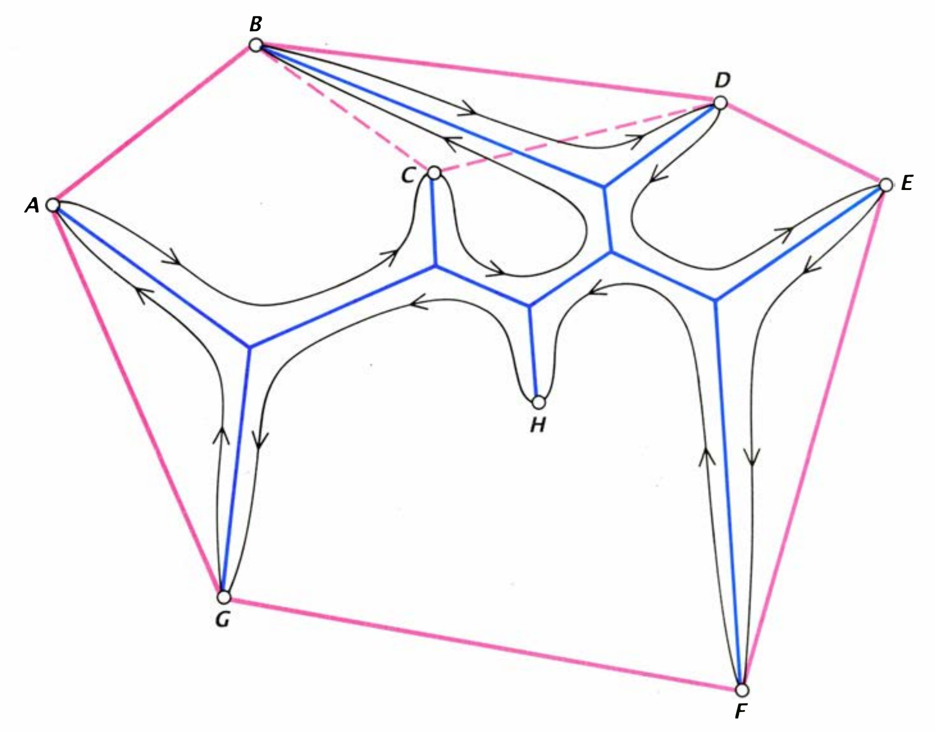
4. EL ALGORITMO DE MELZAK reduce un problema de red mínima a problemas menores. El punto *A* es el lugar correcto para dividir el problema en un problema tripuntual más otro pentapuntual. Para construir posibles árboles de Steiner en el problema pentapuntual, podemos reemplazar un par de puntos (*B* y *C*, por ejemplo) por un único punto (*X* en este caso), construyendo un triángulo equilátero a un lado de *B* y *C*. El problema queda así reducido a cuatro puntos: *X*, *D*, *G* y *A*. A continuación podemos substituir un par de estos puntos: en este caso, primero *D* y *X* con *Y* y después *G* y *A* con *Z*. Se circunscribe una circunferencia a cada uno de los triángulos equiláteros resultantes (*XDY*, *AGZ* y *BCX*). Los puntos en los que la recta que une *Y* con *Z* intercepta a dos de las circunferencias proporcionan los puntos de Steiner *Q* y *R*, y la intersección de la recta que une *X* y *Q* con la circunferencia restante determina el punto de Steiner *P*. Al resultar totalmente inviable determinar por adelantado cuál habrá de ser la mejor repartición y el mejor emparejamiento de los puntos, se hace preciso examinar todas las posibilidades para hallar cuál sea el árbol más corto.

neras de recombinar y formar árboles a partir de los subconjuntos.

Inútil es decir que los investigadores han descubierto métodos superiores para organizar el cálculo y acelerar la velocidad del algoritmo. En vez de considerar la descripción geométrica del problema, se centran en los posibles esquemas de conexiones de la red, en lo que se conoce por topología de la red. La descripción topológica especifica qué puntos se encuentran conectados con otros, pero no precisa las verdaderas ubicaciones de los puntos de Steiner. Dando por supuesta cierta topología, podemos hallar relativamente pronto una secuencia de reemplazo. Tal organización acelera grandemente la velocidad de cálculo de los árboles de Steiner mínimos correspondientes a los subconjuntos. Por ejemplo, para un subconjunto de ocho puntos, el algoritmo necesita considerar sólo unas 10.000 configuraciones topológicas diferentes, en lugar de dos millones de distintas secuencias de reemplazo.

Dado que el número de topologías crece rápidamente con el tamaño del subconjunto, los problemas de Steiner podrían resultar más manejables si tan sólo fuera necesario considerar subconjuntos muy pequeños del conjunto de puntos dados. Experimentos efectuados con el algoritmo de Melzak sugieren que la red más corta correspondiente a seis puntos aleatorios suele ser descomponible en redes mínimas correspondientes a conjuntos de puntos más pequeños. Sin embargo, considerando disposiciones especiales de puntos, denominados "escaleras", Fan R. K. Chung, de Bell Communications Research, y uno de los autores (Graham) demostraron que hay conjuntos de puntos del tamaño que se quiera cuyo árbol de Steiner mínimo no es descomponible. Las escaleras son disposiciones cuyos puntos están situados a intervalos iguales sobre dos rectas paralelas. Se pudo descubrir una solución general para este problema de Steiner tan particular. Dicha solución puso de manifiesto que el número de puntos de Steiner de un árbol mínimo de Steiner correspondiente a una escalera de número impar de "peldaños" es el máximo posible, o sea, el número de puntos dados menos 2. Tal árbol de Steiner no es descomponible, porque la colocación de cada punto de Steiner exige la consideración simultánea de todos los puntos dados. Por tal motivo, no es posible declarar sin más un corte en el tamaño de los subconjuntos considerados en el algoritmo de Melzak.

Cierto número de investigadores mejoraron el algoritmo de Melzak descubriendo métodos más sutiles para reducir la cantidad de trabajo a realizar [*véase la figura 5*]. Tales métodos podan o eliminan partes del cómputo que sólo servirían para producir redes largas. Nuevas técnicas de podado han permitido reducir sustancialmente los tiempos de cálculo. Los programas basados en el algoritmo de Melzak, como el preparado, en 1969, por Ernest J. Cockayne, de la Universidad de Victoria, podían resolver la totalidad de problemas de nueve puntos y ciertos problemas de 12 puntos en cosa de media hora. Un programa reciente, elaborado por Cockayne y Denton E. Hewgill, uno de sus colegas de Victoria, se vale de una poderosa técnica de podadura introducida por Pawel Winter, de la Universidad de Copenhague, para resolver en pocos minutos todos los problemas de 17 puntos y la mayoría de los problemas de 30 puntos engendrados al azar. Tanto éxito tiene el método de podadura de Winter en la eliminación de posibles topologías, que la parte más voluminosa del cómputo consiste ahora en la recombinación de las soluciones correspondientes a los subconjuntos.



5. LOS METODOS DE PODADURA aumentan la eficiencia de los algoritmos para hallar redes pequeñas. Uno de los procedimientos para podar o descartar posibles redes, ideado por Cockayne, consiste en examinar el orden en que una cinta de goma (*en rojo*) tendida en torno al conjunto de puntos dados los va tocando. La banda de goma toca todos los puntos, excepto el *C* y el *H*, aunque *C* puede incluirse en la secuencia, porque el ángulo formado por el punto *C* y dos puntos consecutivos en contacto con la cinta elástica mide al menos 120 grados. El orden de los puntos es entonces *ABCDEFG*. Un circuito cerrado y sin rupturas (*negro*) trazado en torno a una de las posibles redes de conexión (*azul*) toca a los puntos en el orden *ACBDEFHG*. Como *B* y *C* aparecen en orden contrario al establecido por la cinta, esta red hipotética admite ser podada.

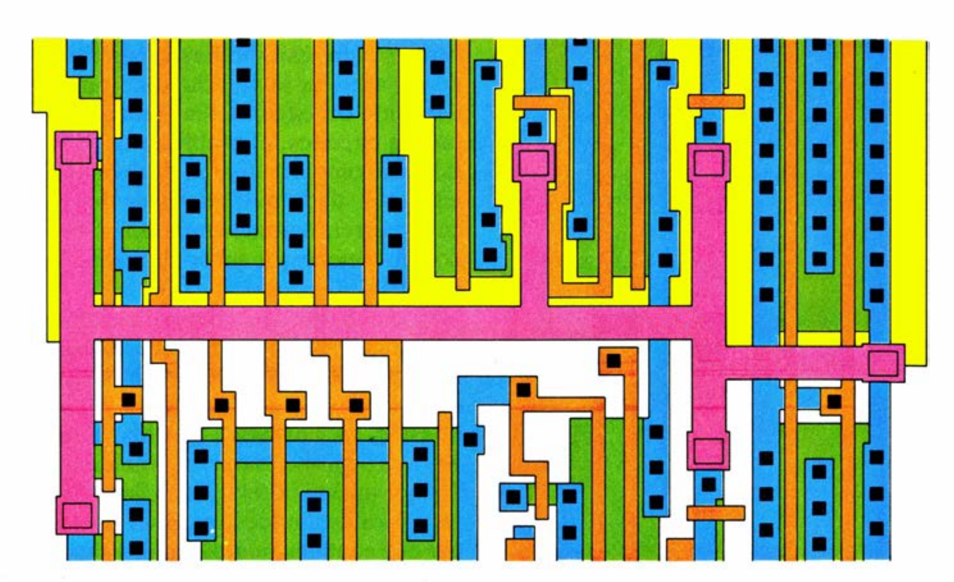
Sin embargo, los tiempos de funcionamiento de cualesquiera de estos programas muestran gran sensibilidad tanto al número de puntos como a la configuración en que éstos se hallen dispuestos. Además, el tiempo de cómputo de los algoritmos, incluso de los más refinados y perfectos, crece exponencialmente con el número de puntos, por lo que los problemas de Steiner de 100 puntos se encuentran hoy muy lejos de nuestro alcance. ¿Llegará algún día a descubrirse un algoritmo eficiente capaz de calcular soluciones de grandes problemas de Steiner?

Los avances en la teoría de computación han convencido a la mayoría de los investigadores de que los algoritmos existentes para los problemas de Steiner no podrán ser mejorados de forma sustancial. Esta teoría asigna un tamaño a cada ejemplo o problema particular. En el caso de los problemas de Steiner, se dispone de un método natural de medición de su tamaño: el número de puntos dados. Se considera entonces el número de operaciones básicas de ordenador —por ejemplo, adiciones, sustracciones y comparaciones— que un algoritmo requiere para

poder resolver un caso particular de determinado tamaño. Dado que dos ejemplos diferentes de un mismo tamaño pueden exigir distinto número de operaciones, se examina cuál es el número máximo de operaciones a realizar en función del tamaño. Si el número de operaciones corresponde a cierta potencia del tamaño (n) del problema, como en el caso de las expresiones n^2, $5n$ o $6n + n^{10}$, se dice que el algoritmo es de tiempo polinómico. Tales algoritmos se consideran eficientes, al menos en un sentido teórico. Cuando el número de operaciones crece exponencialmente con el tiempo, como en 2^n, 5^n o en $3n^2 \times 4^n$, se dice que el algoritmo es de tiempo exponencial.

Aunque en el caso de problemas muy pequeños puedan resultar prácticos tanto los algoritmos polinómicos como los de tiempo exponencial, cuando se trata de problemas grandes los tiempos de solución de los algoritmos de tiempo exponencial resultan tan lentos que tales algoritmos son desesperadamente impracticables [véase "La eficiencia de los algoritmos", por Harry R. Lewis y Christos H. Papadimitriou; INVESTIGACIÓN Y CIENCIA, marzo de 1978]. En el caso de problemas suficientemente grandes, un algoritmo de tiempo polinómico, ejecutado incluso en una máquina muy lenta, produciría una respuesta antes que un algoritmo de tiempo exponencial funcionando en un superordenador.

Si bien se han hallado algoritmos de tiempo exponencial para el problema de Steiner (el algoritmo de Melzak, por ejemplo), no se ha descubierto ningún algoritmo de tiempo polinómico. Y las perspectivas de que llegue a descubrirse un algoritmo eficiente no son buenas. En 1971, Stephen A. Cook, de la Universidad de Toronto, demostró que, de poder hallarse un algoritmo de tiempo polinómico para cualquier problema individual de un grupo de ellos, hoy conocidos por problemas de dificultad NP, tal algoritmo podría emplearse para resolver eficientemente todos los demás problemas de una categoría de problemas difíciles que contiene a los de dificultad NP. Posteriormente, uno de los autores (Graham), en colaboración con Michael R. Garey y David S. Johnson, de los Laboratorios AT&T Bell, demostró que el problema de Steiner era un problema de dificultad NP. Dado que los problemas NP han desbaratado, hasta la fecha, todos los esfuerzos de millares de investigadores, se considera improbable que



6. DETERMINADAS VARIANTES DEL PROBLEMA DE LA RED MINIMA han encontrado aplicación práctica en el diseño de circuitos electrónicos, buscando aumentar su velocidad de funcionamiento. Vemos en color lila la red mínima de hilos verticales y horizontales que interconectan un conjunto de terminales. El fondo de la figura nos muestra otros hilos y terminales situados en capas más profundas.

ninguno de los problemas de dificultad NP, incluido el problema de Steiner, pueda ser resuelto por un algoritmo de tiempo polinómico. No obstante, la demostración de que será imposible resolver eficientemente los problemas de dificultad NP ocupa hoy un lugar sobresaliente entre los problemas de la informática teórica.

Aunque no parece verosímil que se descubra un algoritmo eficiente de tipo polinómico para la determinación por ordenador de las redes mínimas, sí existen métodos prácticos para obtener redes ligeramente más largas. Tenemos un ejemplo en el algoritmo para resolver el más breve árbol de conexión, o árbol de mínimo tendido, que busca la red mínima de segmentos rectilíneos capaz de interconectar un conjunto de puntos dados sin añadir puntos nuevos. Para resolverlo, se conectan los dos puntos dados que se encuentren a distancia mínima, y en cada una de las etapas sucesivas se van conectando el par de puntos más cercanos que se puedan unir sin que al hacerlo aparezca un camino cerrado. Después de todo, si se elimina uno de los lados de un camino cerrado, los puntos dados siguen estando conectados a través de los lados restantes.

Edgar N. Gilbert y Henry O. Pollack, de los Laboratorios Bell, han conjeturado que la razón de un árbol de Steiner mínimo al árbol de conexión mínimo es por lo menos de $\sqrt{3}/2$; es decir, la longitud del árbol de Steiner es a lo sumo un 13,4 por ciento menor que la del árbol de tendido mínimo. La razón $\sqrt{3}/2$ se da en un único ejemplo: tres puntos que formen triángulo equilátero. Aunque tal conjetura está pendiente de demostración, Chung y uno de los autores (Graham) han demostrado que el árbol de Steiner es a lo sumo un 17,6 por ciento más corto que el árbol de tendido.

Con frecuencia resulta posible acortar en un 3 o un 4 por ciento los árboles de tendido minimal añadiendo cuidadosamente puntos de Steiner y reajustando el árbol. Uno de los autores (Bern) ha demostrado que esta clase de algoritmo inexacto posee cierta justificación teórica, dado que la longitud media de un árbol ajustado será algo inferior que la longitud media de un árbol de tendido mínimo.

El árbol de tendido mínimo y los problemas de red mínima han encontrado aplicación en la construcción de sistemas telefónicos, de oleoductos y de redes viarias. Las soluciones, tanto las aproximadas como las exactas, pueden proporcionar directrices para configuración de las redes y la estimación de los volúmenes de materiales necesarios. En versiones más complicadas del problema de Steiner, se puede dar acomodo a la necesidad de evitar ciertos accidentes geográficos o de hallar los enlaces más breves sobre redes preexistentes.

Es posible que la más práctica de las aplicaciones del problema de Steiner se encuentre en el diseño de circuitos electrónicos. En un circuito integrado, una red de líneas de conexión corta exige menos tiempo para cargarse y des-

cargarse que una red larga, permitiendo así acelerar la velocidad de funcionamiento del circuito. Sin embargo, sobre los circuitos, el problema de la red mínima comporta un tipo de geometría diferente, dado que las líneas del circuito corren normalmente tan sólo en dos direcciones; a saber: horizontal y vertical.

El problema, conocido por problema de Steiner rectilíneo, fue abordado en 1965 por Maurice Hanan, del Centro de Investigaciones Thomas J. Watson de IBM, en Yorktown Heights, Nueva York. Lo mismo que en el problema de Steiner inicial, la solución de la versión rectilínea es también un árbol de Steiner que contiene puntos dados y puntos de Steiner, pero ahora los lados concurren formando ángulos de 90 o de 180 grados. Aunque sería concebible que en el problema rectilíneo los puntos de Steiner se encontrasen en un lugar cualquiera, Hanan demostró que en una red rectilínea de longitud mínima resulta posible restringir las ubicaciones de los puntos de Steiner. Se traza, por cada punto, una recta horizontal y una vertical; cada intersección de dos líneas define un posible punto de Steiner. Un algoritmo puede tantear todos los subconjuntos de posibles puntos de Steiner con el fin de calcular una red mínima. Sin embargo, conforme aumenta el número de puntos dados, el tiempo de solución de tal algoritmo "de fuerza bruta" crece exponencialmente. Algoritmos más refinados (pero también de crecimiento exponencial) permiten resolver problemas de Steiner rectilíneos que tengan alrededor de 40 puntos.

Existe una versión rectilínea del problema del árbol de conexión mínima, que puede ser resuelto eficientemente por el algoritmo que elige en cada paso el enlace de longitud mínima, a menos que el enlace dé lugar a un circuito cerrado. Frank K. Hwang, de los Laboratorios Bell, ha demostrado que la longitud de un árbol de Steiner rectilíneo no rebaja nunca en más de un tercio la longitud de un árbol mínimo de tendido rectilíneo.

La aplicación más sorprendente del problema de Steiner se encuentra en el campo de la filogenia. David Sankoff, de la Universidad de Montreal, y otros investigadores han definido una versión del problema de Steiner con el fin de calcular árboles filogenéticos plausibles. Estos investigadores empiezan por aislar una proteína particular que todos los organismos que desean clasificar tienen en común. Seguidamente, determinan para cada organismo la secuencia de los aminoácidos que componen la proteína y definen un punto en una posición determinada por el número de diferencias entre la correspondiente proteína del organismo y la proteína de otros organismos. Los organismos con secuencias similares se consideran próximos, mientras los provistos de secuencias muy diferentes son tenidos por distantes. En la red mínima asociada a esta configuración de puntos dados en un espacio abstracto, los puntos de Steiner hacen el papel de antepasados más plausibles, en tanto que los lados denotan la relación entre organismo y antepasado que comporta mínimo número de mutaciones. Dado que el problema filogenético de Steiner no es más sencillo que los demás problemas de Steiner, este problema ha tenido más importancia como experimento conceptual que como instrumento práctico de investigación, salvo en sus aplicaciones a grupos restringidos de organismos.

Aunque los conocimientos sobre algoritmos han progresado mucho durante estos últimos años, el problema de la red mínima sigue siendo de una dificultad desconcertante. Podemos enunciar el problema en términos sencillos, mas las soluciones desafían el análisis. Una diminuta variación en la configuración geométrica de un problema puede parecer insignificante y, sin embargo, alterar radicalmente la estructura de la red mínima asociada al problema. Tal sensibilidad hace que hasta los aspectos más periféricos referentes a redes mínimas supongan auténticos retos. El problema de la red mínima seguirá fascinándonos, y desesperándonos, durante los años venideros.

BIBLIOGRAFIA COMPLEMENTARIA

¿QUÉ ES LA MATEMÁTICA? Richard Courant y Herbert E. Robbins. Editorial Aguilar. Madrid, 1967 (5ª edición).

STEINER MINIMAL TREES. E. N. Gilbert y H. O. Pollak en *SIAM Journal on Applied Mathematics*, vol. 16, n.º 1, págs. 1-29; enero de 1968.

COMPANION TO CONCRETE MATHEMATICS. Z. A. Melzak. John Wiley & Sons, Inc., 1973.

AN ALGORITHM FOR THE STEINER PROBLEM IN THE EUCLIDEAN PLANE. Pawel Winter en *Networks*, vol. 15, n.º 3, págs. 323-345; otoño, 1985.

STEINER PROBLEM IN NETWORKS: A SURVEY. Pawel Winter en *Networks*, vol. 17, n.º 2, págs. 129-167; verano, 1987.

André-Marie Ampère

Ampère, primer investigador que cuantificó los efectos magnéticos de la corriente eléctrica, fue también pionero de la filosofía de la ciencia. Sus concepciones filosóficas configuraron su método de invención científica

L. Pearce Williams

Cuando mis alumnos me preguntan quién fue André-Marie Ampère, suelo responder, haciendo alusión a los fusibles domésticos: "Posiblemente le conozcáis mejor por el nombre común de ampère; se presenta en dos tamaños, de 15 y 35". Ampère debe su mayor gloria a la creación de la electrodinámica. A comienzos del siglo XIX, realizó las primeras investigaciones sistemáticas sobre los campos magnéticos producidos por corrientes eléctricas, descubrió y cuantificó las fuerzas que actuaban entre hilos conductores recorridos por una corriente y propuso que el magnetismo observado en los imanes se debía a corrientes eléctricas muy débiles que circulaban entre las moléculas de la sustancia magnética. Logros que le valieron que se otorgase su nombre a la unidad internacional de corriente eléctrica: el ampère (A) —unidad en que se miden los fusibles de uso habitual en las viviendas domésticas, de 15 y 35 ampère.

Menos famosos son sus trabajos en el campo de la filosofía de la ciencia, de innegable interés en algunos aspectos. De los científicos máximos de su época, fue el primero en recurrir a las ideas de Immanuel Kant para proporcionar al estudio de la física y la química una fundamentación filosófica. Para Ampère, la teoría del conocimiento de Kant (su teoría acerca de qué pueden saber los seres humanos y cómo llegan a conocerlo) constituyó un punto de arranque para establecer un método práctico de descubrimiento científico. La filosofía kantiana impregnaría las ciencias físicas, especialmente a finales del siglo XIX y principios del XX; el método de descubrimiento, o invención, de Ampère, que inspiró sus trabajos sobre la electrodinámica, perdura todavía en una versión modificada, entre los enfoques científicos más aceptados.

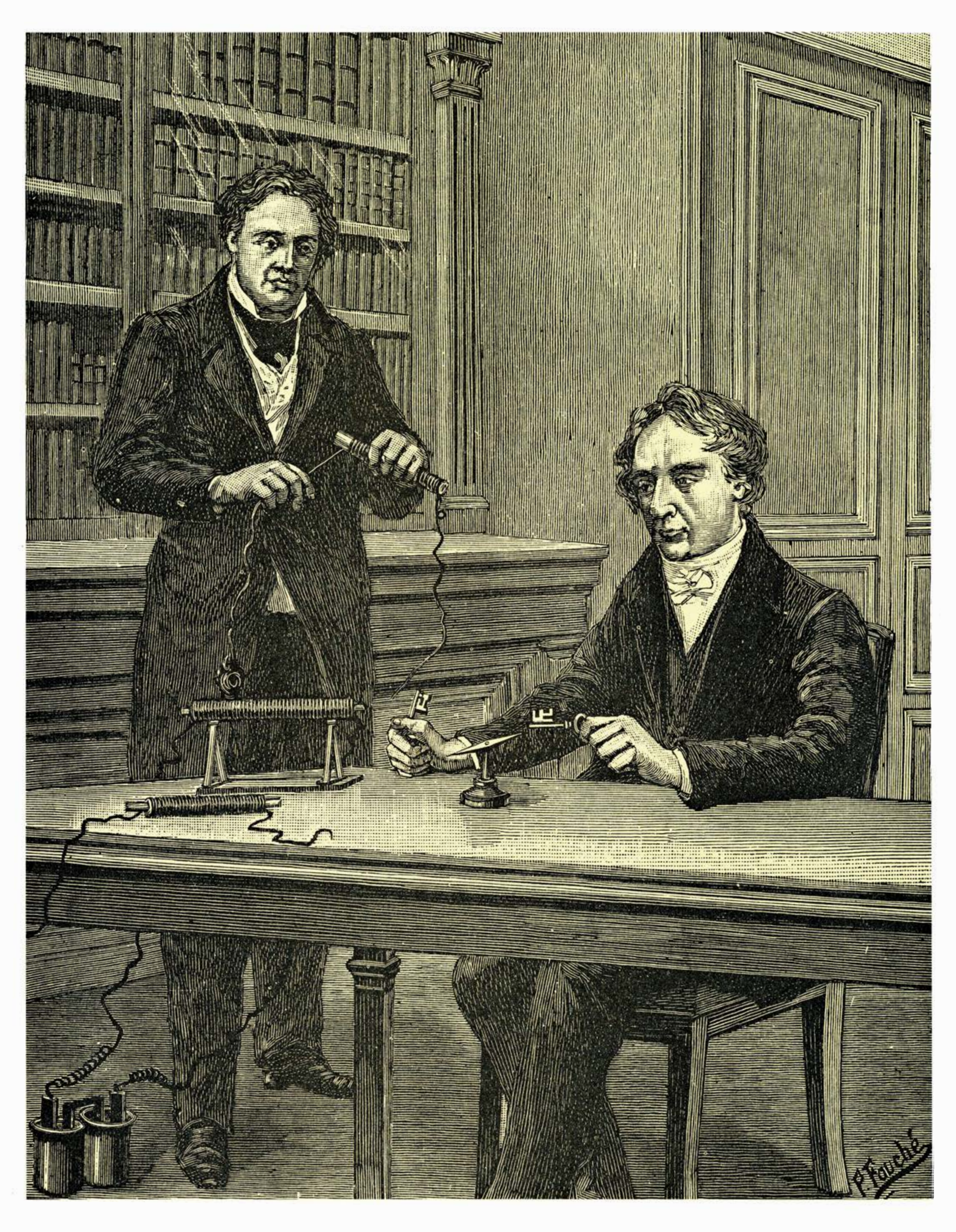
L. PEARCE WILLIAMS es catedrático de historia de la ciencia y director del programa de historia y filosofía de la ciencia y la tecnología de la Universidad de Cornell. Ingresó en esa universidad en 1944. Tras su servicio militar en la Marina, obtuvo su doctorado en 1952. Fue profesor de las universidades de Yale y de Delaware, antes de ingresar en el claustro docente de Cornell. Williams escribe de sí mismo: "Me fascinó la figura de Ampère cuando estaba elaborando una biografía de Michael Faraday. Sus vidas, sus métodos de investigación y sus ideas fueron tan distintas que pensé que resultaría divertido trazar una biografía de Ampère (sobre la que estoy trabajando en la actualidad)".

Nacido en 1775, Ampère se crió en Poleymieux, pedanía de Lyon. Autodidacta, leía con pasión los libros de la biblioteca de su padre y los traídos de la ciudad. Devoró la recién publicada *Encyclopédie* de Denis Diderot y aprendió de memoria artículos enteros, gracias a su memoria fotográfica. Le interesó especialmente la matemática; se abrió camino como pudo en los tratados de álgebra, teoría de las probabilidades y el cálculo del matemático suizo Leonhard Euler. Cuando todavía no contaba veinte años, dominaba ya el innovador libro de Joseph-Louis Lagrange dedicado a la mecánica analítica. Sus intereses más allá de la matemática fueron literalmente enciclopédicos. Ampère estudió los trabajos de historia natural de Georges de Buffon, aprendió griego, latín e italiano, intentó desarrollar un lenguaje universal basado en los principios renovados de la lingüística, se sumergió en la literatura francesa, escribió poesía, estudió botánica y estableció diversos procedimientos para sistematizar sus propias observaciones del mundo natural. Durante toda su vida mantuvo fresco ese afán universalista.

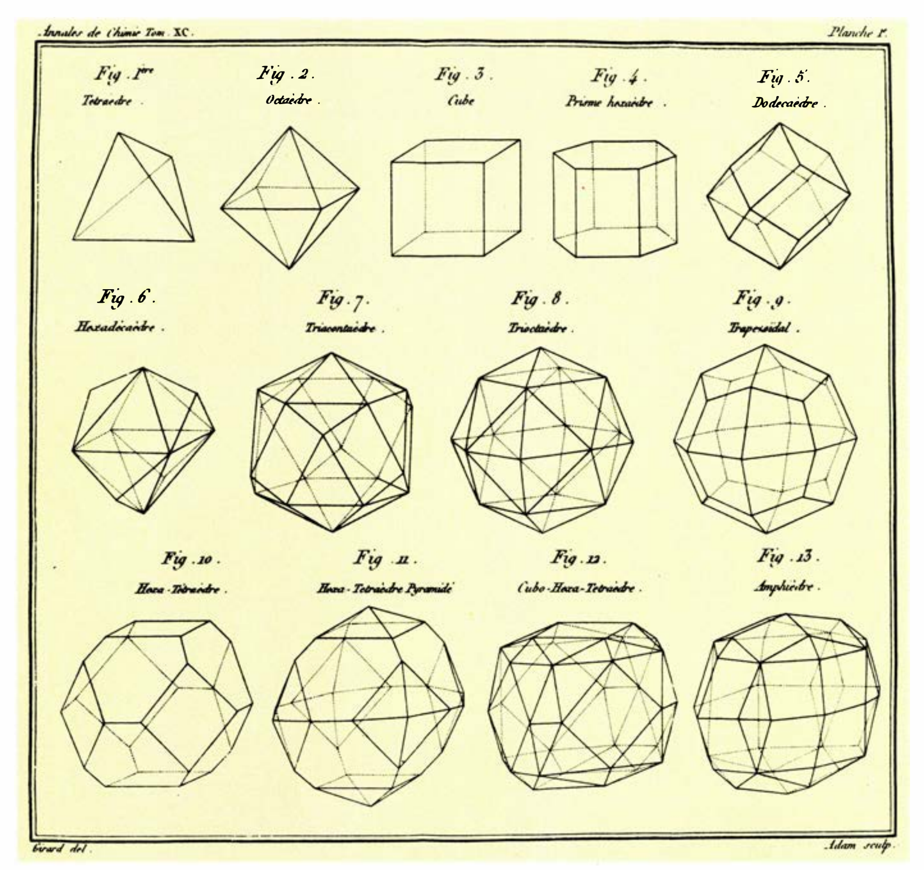
Hasta los 18 años, Ampère, hijo único de una familia acomodada, llevó una existencia casi idílica, que le permitió vagar —física e intelectualmente— por donde le apetecía, en un mundo estable y seguro. En 1793, sin embargo, empezó a desgranar un rosario de tragedias personales que iban a marcar el resto de su vida. Ese año moría su padre en la guillotina, tildado de "contrarevolucionario". En 1803 falleció su esposa, con quien se había casado cuatro años antes y a quien amaba intensamente, madre de su hijo Jean-Jacques. Cuatro años más tarde volvió a casarse y tuvo una hija, pero su matrimonio resultó un fracaso y acabó en divorcio. Tras unos años de calma relativa, su hijo abandonó en 1819 todos sus proyectos y se unió al séquito de Mme. Récamier, gran belleza de la época napoleónica, que coleccionaba admiradores, pero los mantenía a una distancia prudente. Jean-Jacques permaneció en esa situación durante 20 años, sin atender las súplicas de su padre para que volviera a casa y rehiciera su vida. En 1827, la hija de Ampère se casó con un antiguo oficial del ejército de Napoleón, quien resultó ser loco y alcohólico.

A lo largo de su vida, Ampère sufrió un deterioro progresivo de su salud que le impedía, cada vez más, proseguir sus investigaciones científicas. Tuvo también que soportar una situación económica precaria durante la mayor parte de su vida y se vio obligado a aceptar diversos trabajos malpagados, en su mayoría relacionados con la enseñanza de la matemática en distintos grados.

El quehacer científico de Ampère no sólo se resintió por causa de tales circunstancias, sino también por la amplitud de su propia actividad intelectual. En una colección de cartas descubierta recientemente, escritas a la edad de 20 años, vemos cómo salta de la mecánica teórica a la construcción de maquina-



1. ANDRE-MARIE AMPERE (*izquierda*) y su amigo François Arago examinan los efectos magnéticos de las corrientes eléctricas, según se aprecia en esta imagen un tanto idealizada (y bastante imprecisa) de finales del siglo XIX. Ampère fue uno de los fundadores de la electrodinámica (el estudio de las corrientes eléctricas) y adaptó sus ideas filosóficas a la creación de un marco metodológico para la investigación científica. Arago informó a Ampère y a sus colegas de la Academia de Ciencias sobre el descubrimiento del físico Christian Oersted: la corriente eléctrica desviaba una aguja imantada situada en su proximidad.



2. LAS MOLECULAS POLIEDRICAS constituyen la base conceptual del trabajo inicial de Ampère en el campo de la química teórica. Fiel a la idea de Immanuel Kant de que las verdaderas ciencias debían basarse en principios matemáticos, Ampère intentó explicar la química a través de la geometría. En el esquema de Ampère, las moléculas fundamentales constaban de puntos situados en los vértices de los sólidos regulares o casi regulares de la fila superior del gráfico (procedente de un artículo publicado por Ampère en *Annales de Chimie*). Sólo eran posibles las reacciones entre estas moléculas en la hipótesis de que diesen por resultado sólidos con cierto grado de regularidad y simetría, como los ilustrados aquí en las dos filas inferiores.

ria, de ahí a la teoría del vuelo de los cometas, para pasar a los lenguajes artificiales, a la música, a la astronomía, a la botánica, a los sistemas de clasificación. Fue incapaz de concentrarse en un tema determinado y desarrollar en él toda su capacidad intelectual. Mientras elaboraba febrilmente la electrodinámica, siguió especulando sobre problemas de metafísica y filosofía. Nunca vio en ello contradicción alguna, pues estaba convencido de que todo el conocimiento se sustenta en una unidad básica. Como veremos, su último trabajo consistió precisamente en un intento de demostrar esa idea.

Cuando Ampère inició su actividad científica, la filosofía francesa estaba dominada por una escuela cuyos miembros habían sido apodados despectivamente "Idéologues" por Napoleón. Afirmaban éstos haber determinado todas las reglas que rigen el método científico. La mente humana es, decían, un receptor pasivo de impresiones sensoriales y, a partir de ellas, crea una serie de imágenes (que incluyen las sensaciones del olfato, el oído, el gusto y el tacto, además de la vista) que representan el mundo exterior. La mente puede recuperar imágenes y compararlas entre sí para determinar las diferencias y averiguar cómo evoluciona con el tiempo la sucesión de las mismas. Las regularidades observadas pueden utilizarse entonces para fundamentar leyes científicas. No hay forma de probar si existe realmente un mundo exterior. La única realidad conocida está en las imágenes.

En esta visión del mundo queda eliminada la causalidad, reducida a mera sucesión de imágenes. En cierto sentido, resulta imposible, por tanto, explicar los fenómenos y hablar de las fuerzas físicas en virtud de las cuales una causa determinada produce un efecto dado. Siguiendo el razonamiento, no cabe hablar de teorías científicas, en el sentido moderno del término. Por un lado, la ciencia se transforma en taxonomía: el científico distribuye sensaciones análogas en grupos adecuadamente organizados. Por otro, la ciencia adquiere un carácter positivista: el científico expresa regularidades observadas en forma de leyes matemáticamente rigurosas. Un magnífico ejemplo del primer tipo de ciencia lo constituye el sistema químico de Antoine Lavoisier, basado en una descripción exacta, una nomenclatura precisa y una clasificación adecuada de los elementos y compuestos químicos. El segundo tipo de ciencia es el de Jean-Baptiste Fourier en sus estudios sobre el calor; prescindió de las causas del calor y se preocupó por determinar las leyes matemáticas de la propagación del mismo. Ese marco filosófico resulta idóneo para muchos tipos de estudios científicos, pero excluye campos tales como la microfísica, basada en entidades postuladas teóricamente e inobservables directamente: los átomos y las moléculas.

El trabajo científico inicial de Ampère no se sintió aherrojado en ese marco. Empezó su carrera como matemático, y las matemáticas no tienen por qué hacer referencia a la realidad física exterior. Ampère publicó, en 1802, su primer artículo, que versaba sobre la teoría matemática de los juegos de azar, y le valió el renombre necesario para ser contratado como profesor en el liceo napoleónico de Lyon. En ese opúsculo, demostraba que un jugador que dispusiera de una cantidad finita de dinero, frente a otro jugador con capacidad económica infinita o a varios jugadores con recursos finitos, lo perdería todo al cabo de un tiempo finito. Escribió también varios artículos relacionados con la mecánica teórica, pero su obra matemática más extensa —que escribió para ocupar un sillón en la Academia Francesa de Ciencias— consistió en un tratado sobre ecuaciones en derivadas parciales. En esta fase de su carrera escribió otros artículos matemáticos; si se hubiese dedicado exclusivamente a este dominio del saber, hoy se le consideraría, en el mejor de los casos, un matemático competente y en ocasiones innovador, pero quedaría eclipsado por sus contemporáneos Laplace, Poisson, Cauchy y Fourier.

Si fracasó a la hora de abrir nuevas perspectivas en el campo de las matemáticas debióse a que esta disciplina llegó a aburrirle. Ocurría eso en 1805; se había apasionado por dos nuevos temas: la metafísica y la química.

La fascinación de Ampère por la metafísica surgió con motivo de su encuentro con los "Idéologues" en 1804, cuando abandonó Lyon para vivir en

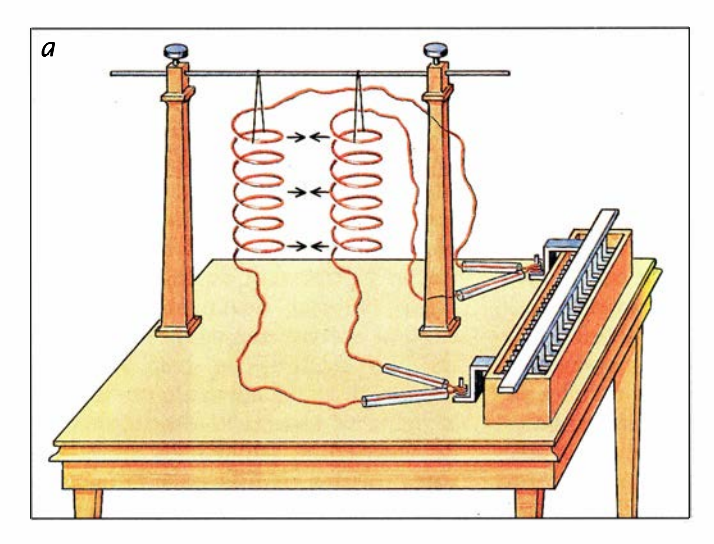
París, tras la muerte de su primera esposa. Ampère frecuentó un pequeño grupo de filósofos que se reunían en Auteuil, en los aledaños de París, para discutir sobre la obra del fundador de la escuela, Condillac. No tardó en rechazar los puntos de vista de aquella filosofía que negaba la existencia de Dios y del alma inmortal. Se separó del grupo, buscando una alternativa. Encontró la inspiración de su propia filosofía en los escritos de Kant.

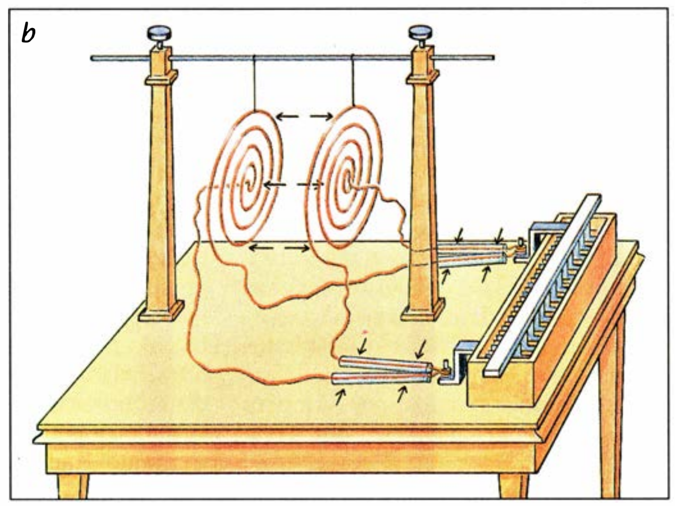
Kant dividía el mundo en dos ámbitos: el de los fenómenos y el de los noúmenos. Los fenómenos son acontecimientos percibidos por la mente humana; son las sensaciones. Los noúmenos son las causas de los fenómenos; son las llamadas "cosas en sí", los objetos realmente existentes. Los seres humanos no pueden conocer directamente los noúmenos. Estos son las fuentes de las señales que actúan sobre nuestros sentidos; sólo podemos percibir las señales, pero no las fuentes. Por tanto, según Kant, nunca podremos saber nada definitivo sobre los noúmenos.

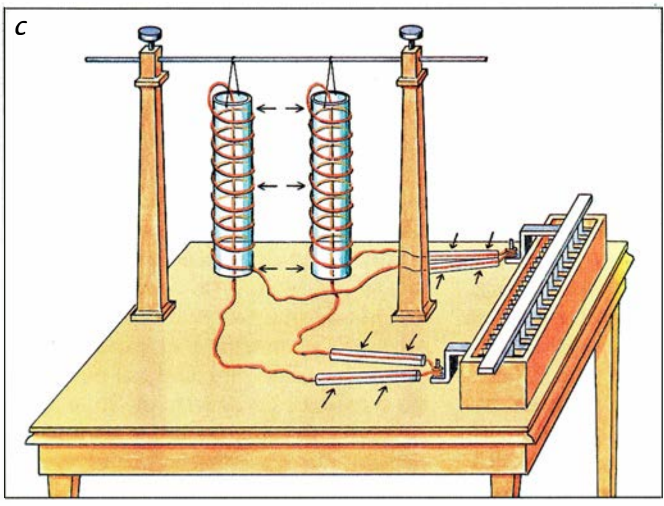
En este punto, Ampère se separaba de Kant. Advirtió aquél que, entre los fenómenos, se producían a menudo vínculos, *rapports*, según su propia expresión. *Rapports* que tenían que ser análogos a los que se daban entre los noúmenos subyacentes a los fenómenos. Por consiguiente, conoceríamos las interacciones entre los noúmenos inobservables si estudiásemos los *rapports* entre los fenómenos.

Sobre esa proposición filosófica asentó Ampère su método de invención científica. Lo expuso en una carta escrita, en 1810, a su viejo amigo Marie-François Pierre Maine de Biran. En el método de Ampère, el científico explica los fenómenos mediante hipótesis sobre la existencia de ciertos noúmenos. Emprende luego un proceso de deducción: una vez aceptada la existencia de entidades teóricas, ¿qué nuevos resultados experimentales, qué nuevos fenómenos cabe esperar? La deducción requiere la comprobación del experimento. Se puede suponer la existencia de entidades teóricas mientras proporcionen una explicación comprobable de los fenómenos. La probabilidad de que la teoría sea verdadera aumenta conforme pone de manifiesto su capacidad para resistir la evidencia experimental. El de Ampère suele llamarse hoy método hipotético-deductivo y son muchos los epistemólogos que lo consideran la forma adecuada de conducir la investigación científica.

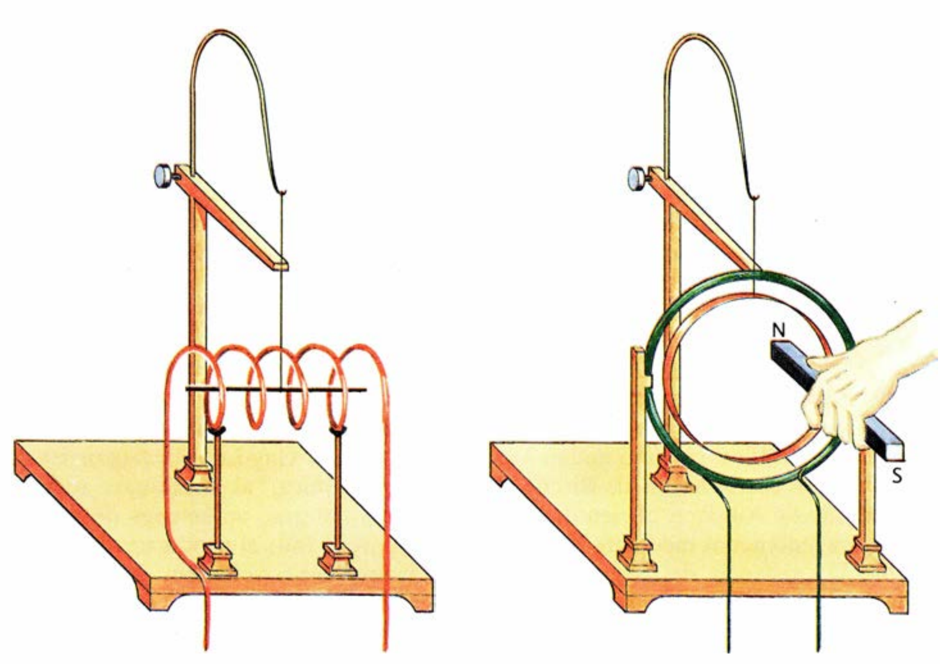
Un buen ejemplo de aplicación del método se encuentra en el análisis de Ampère sobre la ley de combinación de volúmenes postulada por Joseph-Louis Gay-Lussac. Según esta ley fenoménica, al combinar, según números enteros, volúmenes de gases en una reacción, el producto resultante ocupa otro volumen, expresable también por un número entero; por ejemplo, cuando se combinan dos litros de hidrógeno con un litro de oxígeno, el resultado ocupa precisamente dos litros de vapor de agua. Para los "Idéologues", no se podía ir más allá; no se podía justificar la certeza de esa ley. Ampère sí estaba resuelto a dar un paso más. En 1814, en uno de los







3. SERIE DE EXPERIMENTOS realizados para comprobar la hipótesis fundamental de Ampère según la cual el magnetismo (incluido el de los imanes permanentes) está provocado por corrientes eléctricas que recorren trayectorias circulares. En opinión de Ampère, si la hipótesis fuese cierta, entonces las hélices por las que circula corriente en el mismo sentido (*a*) se repelerían entre sí, a modo de imanes permanentes cuyos polos norte apuntasen en la misma dirección. Las hélices, por contra, se atraerían. Cuando Ampère realizó este experimento con espirales planas, las espirales se comportaron como imanes (*b*). Durante el experimento, Ampère advirtió que, cuando se hacía pasar corriente en el mismo sentido por unos hilos rectos próximos a la batería (*derecha*), éstos se atraían entre sí. Concluyó que, al haber utilizado hélices muy poco densas, el flujo circular de corriente había sido insignificante con respecto al flujo longitudinal por la hélice y, así, las hélices se habían atraído mutuamente, como lo hubieran hecho dos alambres rectos. Comprobó su hipótesis mediante hélices enrolladas sobre tubos de vidrio, por los que hizo pasar un extremo del hilo, con objeto de que la corriente de regreso pudiese contrarrestar los efectos de la corriente longitudinal. Entre hélices así arrolladas, como se muestra en una versión simplificada del experimento (*abajo, a la izquierda*), se produjo una repulsión. (La reconstrucción de los ensayos ha sido realizada por Hank Iken.)



4. NATURALEZA DE LAS CORRIENTES que provocan el magnetismo. De ella se ocuparon, en sendos experimentos, Arago y Ampère. En el primero (*izquierda*), Arago utilizó un hilo de cobre enrollado en forma de hélice, colocó una aguja de hierro en su centro e hizo circular corriente por la hélice. La aguja quedó imantada. Ampère se propuso comprobar si las corrientes circulares que originaban el magnetismo permanente circulaban alrededor de cada partícula de la sustancia imantada o alrededor del eje del imán considerado globalmente. En su experimento (*derecha*), colocó una bobina de hilo de cobre aislado (*verde*) y suspendió en su interior un anillo de cobre. Si la bobina hubiera imantado la aguja creando corrientes que circulasen alrededor de su eje central, pensaba Ampère, también tendría que haber imantado el anillo de cobre, pues las corrientes podrían circular alrededor de su eje. Hizo pasar la corriente por la bobina; para ver si el anillo se había convertido o no en un imán ''permanente'', colocó un imán alargado cerca de aquél. Al comprobar que no se había convertido en un imán permanente, Ampère se convenció de que las corrientes de los imanes permanentes circulan alrededor de las distintas partículas componentes, y no del eje magnético. El experimento, aunque no la conclusión a que llegó, es defectuoso, visto desde una óptica moderna.

primeros artículos modernos de química teórica, Ampère afirmaba que sólo se podía explicar el fenómeno suponiendo que volúmenes iguales de distintos gases a la misma temperatura y la misma presión debían contener el mismo número de moléculas. La existencia de las moléculas —noúmenos inobservables— constituye la base de la explicación dada por Ampère a los fenómenos descritos por Gay-Lussac.

Poco después de idear su método, Ampère lo aplicó a la creación de un brillante marco teórico para la química. Kant había insistido en el hecho de que las ciencias debían fundamentarse en principios matemáticos a priori, principios que creía inherentes a la estructura de la mente humana y no basados en la observación. A partir de esa idea de Kant, Ampère intentó deducir las leyes de la afinidad química (las que determinan qué reacciones son posibles), suponiendo la existencia de hipotéticas moléculas de naturaleza geométrica.

Ampère imaginó que cada molécula estaba constituida por átomos puntiformes distribuidos en el espacio cual vértices de un sólido geométrico sencillo, como el tetraedro, el octaedro o el cubo. Las únicas reacciones químicas posibles son aquellas cuyo resultado es un sólido geométrico que posea cierto grado de regularidad y simetría tridimensional. En el marco teórico establecido por Ampère, la enigmática arbitrariedad de la actividad química podía reducirse a una certeza matemática: la química podía basarse en la geometría, que Kant consideraba la forma más pura de la matemática. Hay que señalar que ninguno de los artículos de Ampère dedicados a la química suscitó adhesión alguna por parte de los demás químicos, quienes se mostraron hostiles tanto a sus especulaciones como a sus matemáticas.

Lo realizado por Ampère hasta 1819, cuando tenía 44 años, apenas si ocuparía una breve nota a pie de página de cualquier historia de la física. A diferencia de sus contemporáneos Augustin Fresnel (gran amigo suyo y creador de la teoría ondulatoria de la luz) y Sadi Carnot (uno de los fundadores de la termodinámica), quienes murieron, respectivamente, a los 39 y 36 años, Ampère desarrolló su principal trabajo científico en plena madurez, cuando ya desesperaba de hacer alguna aportación crucial.

En sus primeros experimentos sobre electrodinámica, Ampère utilizó una pila voltaica, como la que había inventado Alessandro Volta en 1800. La pila voltaica es un sistema electroquímico, con algún parecido con la batería de un automóvil actual. Al conectar mediante un hilo conductor los polos de una pila voltaica, se produce una circulación de corriente por el hilo y las reacciones químicas que tienen lugar en el interior de la pila mantienen la diferencia de potencial entre los polos.

La corriente eléctrica continua constituía un fenómeno novedoso a comienzos del siglo XIX; las primeras teorías sobre ella se basaron en las de la electrostática. Pocos pensaron que la corriente ejerciera efectos magnéticos, pues Charles de Coulomb había demostrado, en la década de 1780, que las fuerzas asociadas a la electrostática diferían de las del magnetismo. Tan sólo algunos "filósofos de la naturaleza", convencidos de la unidad de todas las fuerzas, buscaron una relación entre la electricidad y el magnetismo.

Uno de estos filósofos, de origen danés, se llamaba Hans Christian Oersted. Había publicado, en 1807 y 1812, diversos trabajos en los que insistía, desde concepciones filosóficas, que debía existir una relación entre la electricidad y el magnetismo. En el invierno y primavera de 1820, había logrado detectar finalmente una relación, al acercar una aguja imantada a un hilo eléctrico de gran longitud; cuando circulaba corriente por el hilo, la aguja se movía. El descubrimiento de Oersted se publicó en las principales revistas científicas de la época.

La noticia llegó a París por boca de François Arago, amigo de Ampère y quien había presenciado ese efecto durante una visita a Ginebra. Los miembros de la Academia de Ciencias escucharon con escepticismo el informe presentado por Arago y sólo quedaron convencidos tras la demostración que éste hizo el 11 de septiembre de ese mismo año. Ampère asistió a la demostración y decidió estudiar ese efecto a su regreso a casa. Se dio cuenta en seguida de que Oersted no había comprendido plenamente el experimento: no había tenido en cuenta el magnetismo terrestre. La desviación de la aguja dependía del ángulo entre el hilo recorrido por la corriente y el campo magnético terrestre.

Ampère se puso, de inmediato, a buscar el verdadero efecto de la corriente eléctrica sobre una aguja imantada. Colocó una serie de imanes en rotación libre de manera que el campo magnético terrestre quedase neutralizado en una pequeña región. Con gran satisfacción, comprobó que la aguja se desviaba ahora hasta formar un ángulo recto con el hilo recorrido por la corriente. Se dio cuenta entonces de que la aguja imantada podía ser un elemento de instrumentos que permitiesen detectar el paso de una corriente eléctrica. Con este nuevo instrumento, que llamó galvanómetro, determinó la intensidad de la corriente a lo largo de un circuito compuesto por un hilo conductor y una pila voltaica.

Se venía suponiendo, hasta entonces, que los mecanismos que operaban en el interior de una pila voltaica diferían de los mecanismos de la corriente que circulaba entre los dos polos de la pila. Con gran sorpresa, Ampère comprobó que la intensidad de la corriente en la pila voltaica era igual que la del resto del circuito. ¿Qué pasaría entonces, pensó, si se construía una pila circular, es decir, tal que su polo positivo estuviese en contacto con su polo negativo? Existen razones para pensar que, en ese fecundo mes de septiembre de 1820, construyó dicha pila y observó que creaba un campo magnético simétrico. Ampère concibió entonces la hipótesis que iba a defender durante el resto de su vida: el magnetismo no consiste sino en corrientes eléctricas que viajan por trayectorias circulares. Hasta allí había llegado el 18 de septiembre, cuando leyó su primer informe ante la Academia. Nadie le prestó demasiada atención.

Ampère dedicó la semana del 18 al 25 de septiembre, fecha de la siguiente reunión de la Academia, a la aplicación de su método. Después de sentar la hipótesis de una explicación nouménica —corrientes eléctricas circulares— del fenómeno del magnetismo, el siguiente paso consistía en comprobar las deducciones basadas en dicha hipótesis: demostrar que las corrientes eléctricas circulares producían los mismos efectos que los imanes permanentes. Primero intentó mostrar el efecto con unos hilos de cobre que había enrollado formando hélices (solenoides en muelle). Cuando ponía juntas dos hélices y hacía pasar simultáneamente por ambas una corriente en la misma dirección [*véase la figura 3*], las bobinas debían repelerse entre sí, como dos imanes acercados con sus polos norte en la misma dirección. Sin embargo, comprobó, para su sorpresa, que las bobinas se atraían.

Ampère probó entonces otro sistema. Arrolló el hilo de cobre de modo que formase espirales planas, con un extremo del hilo en el centro de la espiral y el otro en el borde de la misma. En esta ocasión, cuando hizo pasar la corriente, manteniendo juntas las espirales, éstas se comportaron como imanes. El 25 de septiembre, mostró a sus colegas los diversos casos de atracciones y repulsiones.

¿Por qué se comportaban de manera distinta las hélices? La respuesta la encontró en una observación fortuita. Los hilos de las dos espirales se hallaban conectados a la misma pila, y así los hilos que estaban conectados a un mismo polo quedaban muy próximos entre sí. Cuando Ampère hizo pasar la corriente, observó que los hilos por donde circulaba corriente en el mismo sentido se atraían mutuamente, aun cuando no estuviesen enrollados en carrete. Esta interacción entre líneas rectas, no observada hasta entonces, suministró a Ampère la respuesta al problema de las hélices. Puesto que no había aislado los hilos, dejó enrolladas las hélices sin apretarlas: las bobinas que constituían cada hélice distaban así bastante. Por consiguiente, las hélices se habían comportado como dos hilos rectos recorridos por la corriente en el mismo sentido y, por tanto, se había producido la mutua atracción.

Para someter a prueba esta conclusión, Ampère enrolló una hélice alrededor de un tubo de vidrio y pasó el extremo del hilo por el centro del mismo. Esperaba que la corriente que circulaba por el centro del tubo cancelase los efectos longitudinales de la corriente entre los dos extremos de la hélice, para apreciar así los efectos de la componente circular de la corriente. Las hélices construidas según ese sistema se comportaron como auténticos imanes permanentes, confirmándose la hipótesis de Ampère.

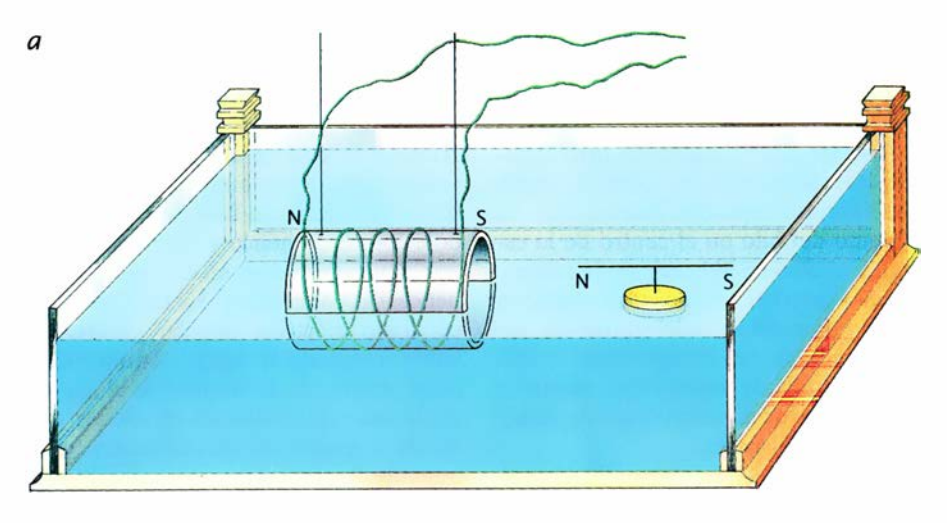
Pero los resultados planteaban una cuestión peliaguda: ¿dónde se ocultaban las corrientes eléctricas en un imán permanente? Sólo había dos posibilidades: las corrientes recorrían círculos alrededor del eje del imán considerado como un todo o recorrían círculos mucho menores alrededor de cada una de las partículas de que constaba el imán. Siguiendo una sugerencia de Fresnel, Ampère supuso que las corrientes circulaban alrededor de las moléculas componentes del imán.

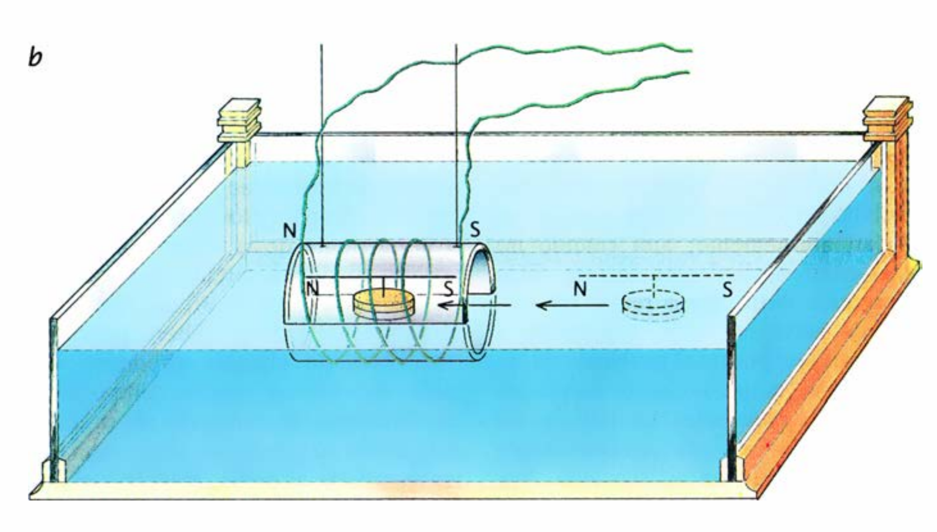
La hipótesis demandaba una comprobación experimental, para la cual Ampère se basó en un experimento llevado a cabo por su amigo Arago, quien había enrollado un hilo de cobre dándole forma de hélice y había colocado una aguja de hierro en el centro de la hélice. Al conectar ésta a la batería, la corriente eléctrica circular resultante había imantado la aguja. Ahora bien, ¿qué ocurría en el interior de la aguja: circulaban las corrientes alrededor del eje de la aguja o de las moléculas componentes?

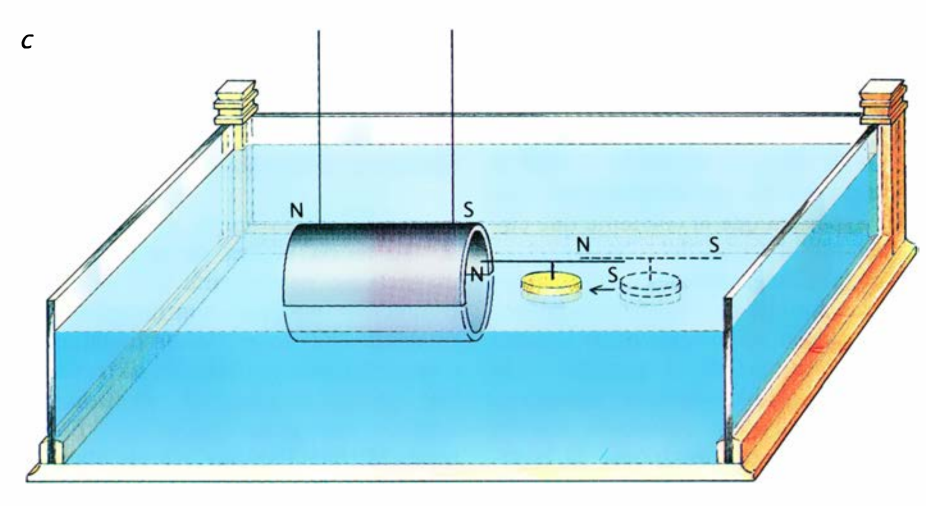
Ampère se aprestó a resolver la cuestión construyendo un fino anillo de cobre que suspendió en el interior de una bobina cilíndrica de hilo de cobre aislado, de manera que el diámetro del anillo fuese ligeramente inferior al de la bobina. Ampère dispuso ambos elementos de forma que fuesen concéntricos y de ejes paralelos. Según Ampère, si la corriente de la hélice había provocado corrientes circulares alrededor del eje de la aguja de Arago, entonces una corriente análoga debería provocar una corriente circular alrededor del anillo de cobre y éste se comportaría transitoriamente como un imán. Para comprobar su previsión, acercó una barra imantada al anillo, mientras circulaba corriente por la bobina. Si el anillo hubiese quedado imantado, la barra imantada hubiese provocado una repulsión, pero, como esperaba Ampère, ésta no se produjo. Repitió en público su experimento, y lo consideró una sólida confirmación de su hipótesis de las corrientes moleculares.

Por esa misma época, sin embargo, su teoría acerca de la causa del magnetismo permanente tuvo que enfrentarse a una grave contradicción. A finales de 1821 se publicó en Inglaterra, e inmediatamente traducida al francés, una historia anónima del electromagnetismo, cuyo autor resultó ser el físico inglés Michael Faraday. En ese libro se explicaban dos experimentos que pretendían rebatir la hipótesis central de Ampère, según la cual el magnetismo permanente era el resultado de corrientes eléctricas circulares. Para Faraday, los datos experimentales demostraban que el magnetismo de los imanes permanentes difería sustancialmente del magnetismo de las hélices recorridas por una corriente.

En el primer experimento intervenía una hélice de hilo de cobre aislado, enrollada alrededor de un tubo de vidrio,







5. CONTRADICCION APARENTE en la hipótesis fundamental del magnetismo de Ampère. Tal se desprendía de los experimentos realizados por Michael Faraday. Faraday enrolló una hélice de hilo de cobre aislado alrededor de un tubo de vidrio hueco y sumergió parcialmente el tubo en una cubeta de agua (*a*). Colocó una larga aguja imantada sobre un corcho que flotase cerca del tubo e hizo pasar la corriente por la hélice. Según Faraday, si la hélice portadora de corriente se comportase como un auténtico imán, el corcho se encaminaría hacia el tubo hasta que la aguja penetrara en el interior de la hélice; uno de los polos magnéticos de la aguja se encontraría entonces lo más cerca posible del polo opuesto de la hélice-imán. No hubo tal: el corcho se desplazó por el interior del tubo hasta que los polos de la aguja se aproximaron a los respectivos polos de la hélice (*b*). En otro experimento (*c*), Faraday colocó una lámina enrollada de acero en el interior del tubo, lo imantó y lo sumergió parcialmente en agua. En este ocasión, como esperaba, la aguja se desplazó hasta un extremo del tubo y quedó inmóvil. Faraday dedujo que la hélice recorrida por la corriente y el imán no eran idénticos. (Ampère logró salvar esta contradicción aparente en su hipótesis advirtiendo que, en la hélice, la aguja se encontraba en el centro de la corriente, mientras que, en el tubo de acero, quedaba al margen de todos los pequeños circuitos por los que pasaba la corriente; por consiguiente, cabía esperar que la aguja se comportase, en una situación, de forma distinta de como lo hiciera en la otra.)

hueco y ancho. Se sumergía parcialmente el tubo en una cubeta con agua, de manera que el eje longitudinal de la hélice fuese paralelo a la superficie del agua. Se colocaba una larga aguja imantada sobre un corcho que flotase en la cubeta, cerca del tubo, y se hacía pasar la corriente por la hélice. Según el razonamiento de Faraday, si la hélice se comportara como un imán genuino, el corcho debería entonces desplazarse hasta el tubo de vidrio y quedarse allí inmóvil, ya que uno de los polos de la aguja (por ejemplo, el polo norte) se habría acercado lo más posible al polo opuesto de la hélice-imán (el polo sur). En lugar de ello, se observaba que el corcho se encaminaba hacia un extremo del tubo para moverse luego a lo largo de éste (por el centro de la hélice) hasta que ambos polos de la aguja quedaban inmóviles bajo los polos análogos (y no opuestos) de la hélice-imán.

En el segundo experimento se ponía de manifiesto el efecto en el caso de un imán permanente auténtico. Se fabricaba un tubo hueco con una lámina de acero enrollada, se magnetizaba y se sumergía parcialmente en el agua. En esta ocasión, el corcho se desplazaba hasta el tubo y quedaba allí inmóvil, sometido al polo sur del imán, lo cual, según Faraday, confirmaba que las hélices recorridas por una corriente no son lo mismo que los imanes permanentes y que, por tanto, el magnetismo permanente no es el resultado de corrientes eléctricas circulares.

Ampère consiguió resolver el problema planteado. Sostuvo que, si en los imanes las corrientes circulan alrededor de las moléculas componentes, entonces el centro de un tubo de acero debía diferir cualitativamente del centro de la hélice. En la hélice, la aguja imantada había quedado "encerrada" en el interior de la corriente en circulación; en el tubo de acero, por contra, la aguja permanecía a resguardo de todas las corrientes moleculares. Por consiguiente, cabía esperar que la aguja se comportase de forma distinta en uno y otro caso. Tras esta explicación (básicamente correcta), Ampère quedó públicamente comprometido con el concepto de corrientes moleculares.

La intensidad del compromiso de Ampère se puso de manifiesto en el verano de 1822, cuando repitió el experimento del anillo y la bobina de cobre. En esta ocasión, utilizó un potente imán en forma de herradura y no una pequeña barra imantada; observó que el anillo resultaba repelido. La reac-

ción de Ampère ante este fenómeno se hace difícil de entender. El experimento parecía contradecir su teoría de las corrientes moleculares, pero él no hizo ninguna comprobación ulterior. Mencionó de pasada su experimento en su informe de septiembre a la Academia y apostilló el sorprendente comentario de que el efecto carecía de significado teórico. De hecho, Ampère había observado, sin darse cuenta, la inducción electromagnética de una corriente sobre otra, pero hasta 1832, cuando Faraday descubrió e investigó la inducción electromagnética, Ampère no se percató de cuán cerca había estado de un descubrimiento tan importante.

A partir de 1822, Ampère ya no modificaría sus conceptos básicos de la electrodinámica, si bien varió su capacidad de cuantificar la teoría. Una vez hubo formulado su teoría del magnetismo permanente, se centró en la determinación experimental de la intensidad de las distintas fuerzas electromagnéticas. Ampère concluyó que la interacción electromagnética fundamental era la fuerza que actuaba entre dos hilos por donde circula una corriente. Se planteó la difícil tarea de medir dicha fuerza. También aquí, su ingenio para plantear hipótesis le sirvió de gran ayuda. A diferencia de las fuerzas gravitatorias, que pueden abordarse, desde un punto de vista matemático, como fuerzas que actúan sobre puntos geométricos, las fuerzas debidas a la corriente no siempre pueden considerarse meros fenómenos locales.

La idea de Ampère consistió en considerar segmentos infinitésimos de los hilos recorridos por la corriente y en suponer que las fuerzas que actuaban entre ellas eran inversamente proporcionales al cuadrado de la distancia que las separaba. Calculó entonces la fuerza total entre los dos hilos integrando —sumando matemáticamente todas las unidades infinitésimas de fuerza— y teniendo en cuenta las direcciones de actuación de las fuerzas en función de la curvatura de los hilos. Ampère consideró primero el caso particular en que los dos hilos se hallaban en el mismo plano. Más tarde, generalizó el resultado a los hilos situados en planos distintos, para cualquier ángulo entre ellos. Esta generalización posibilitó estudiar el caso de alambres retorcidos en cualquier sentido dentro del espacio tridimensional. El resultado fue la fórmula compacta gracias a la cual puede calcularse la fuerza electromagnética que actúa entre dos hilos cuando se conoce la intensidad de las corrientes y la geometría de los cables. En 1826, Ampère fundió todos sus artículos anteriores en un libro definitivo, *Mémoire sur la théorie mathématique des phénomènes électrodynamiques uniquement déduite de l'expérience*.

La salud de Ampère fue deteriorándose precipitadamente desde 1827. Abandonó la investigación científica y se dedicó a la filosofía de la ciencia, recuperando el talante de su primera juventud. Quedó fascinado por la doctrina de la armonía prestablecida de Gottfried Wilhelm von Leibniz, según la cual la mente del ser humano es una copia, si bien imperfecta, de la mente de Dios. Como el proceso de raciocinio del ser humano es una imagen del de Dios, la mente humana debería poder entender el universo a través del razonamiento puro. Dicho de otra forma, debe existir una armonía prestablecida entre las leyes del universo y la capacidad de razonamiento del ser humano.

Ampère llegó a la conclusión de que la correspondencia entre la razón de Dios, la razón humana y la racionalidad inherente del universo debían permitir el recurso a un proceso de taxonomía básicamente para encontrar la Verdad última. Ampère creía que si se consiguiesen explicar todas las ciencias que la mente humana es capaz de construir, se lograría entonces la clave de toda verdad posible, ya que la mente está estructurada de forma que se corresponde directamente con la estructura del universo. Bastaría entonces con llenar de contenido el mapa taxonómico cósmico.

Ampère elaboró diversos mapas de ese tipo que pretendían ser herramientas para la investigación fundamental, antes de morir en 1836, presumiblemente de neumonía. Había acabado casi como empezó; un enciclopedista comprometido con la unidad del conocimiento, pues éste no es más que el reflejo de la unidad de la mente divina.

BIBLIOGRAFIA COMPLEMENTARIA

AMPÈRE. L. PEARCE WILLIAMS en *The Dictionary of Scientific Biography*. Charles Coulston Gillespie (ed.), Charles Scribner's Sons, 1970.

A.-M. AMPÈRE ET LA CREATION DE L'ELECTRODYNAMIQUE (1820-1827). Christine Blondel. Comité des travaux historiques et scientifiques, 1982.

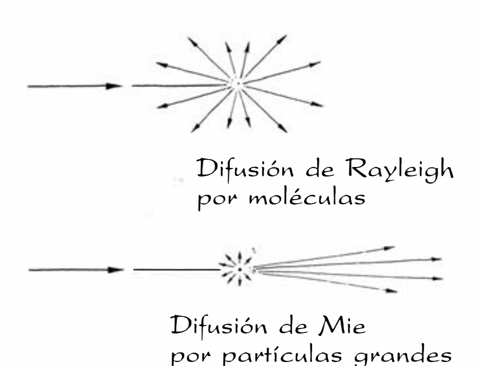
WHAT WERE AMPÈRE'S EARLIEST DISCOVERIES IN ELECTRODYNAMICS? L. Pearce Williams en *Isis*, vol. 74, n.º 274, págs. 492-508; diciembre de 1983.

Taller y laboratorio

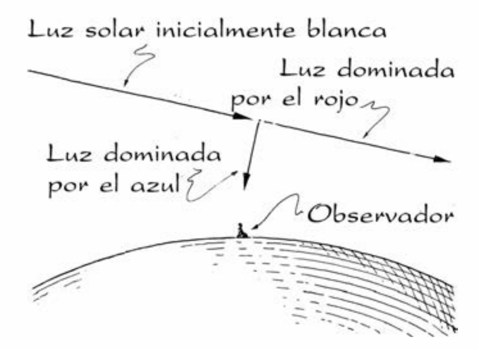
Los colores que observamos en el cielo nos ofrecen lecciones de difusión óptica

Jearl Walker

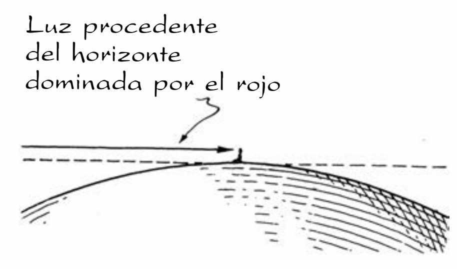
Pese a los concienzudos estudios científicos a que se ha sometido el tema durante más de un siglo, la explicación de los colores que ofrece el cielo durante el día y el crepúsculo ha presentado grandes dificultades. ¿Por qué el firmamento de un día claro se nos muestra azul en su mayor parte, para tornarse blanco junto al horizonte? ¿Por qué el sol del ocaso suele ser rojo y el cielo justo encima suyo parece un tapiz cromático? ¿Por qué, en el crepúsculo, se alza por levante una sombra curva con un borde rosado? ¿Por qué, por poniente, aparece, a veces, una mancha púrpura, que luego se desvanece, poco después del ocaso y por qué, también a veces, aparece otra mancha púrpura, algo así como dos horas más tarde? Las respuestas a estas preguntas precisan del estudio de la interacción entre la luz, por un lado, y las moléculas del aire y las partículas suspendidas en éste, por otro. En algunos casos, se están buscando todavía respuestas definitivas.



1. Dos tipos de difusión



2. Separación de los colores por difusión de Rayleigh



3. Color del sol en el ocaso

Las explicaciones del motivo por el que un cielo claro sea en su mayor parte azul no han escaseado. En los modelos más conocidos, se hace intervenir la difusión de la luz por sustancias en suspensión en el aire, tales como polvo, aerosoles, cristales de hielo y gotitas de agua; en otros, se recurre a la absorción del extremo rojo del espectro visible por parte del agua y el ozono de la atmósfera. Las insuficiencias de tales explicaciones las analizaron, en 1985, Craig F. Bohren y Alistair B. Fraser, de la Universidad de Pennsylvania, quienes expusieron asimismo la explicación correcta, avanzada por Lord Rayleigh en 1899.

El mismo Rayleigh se mostró remiso a aceptar su propia explicación, en parte a causa de los descubrimientos publicados, en 1869, por John Tyndall, físico británico a quien se recuerda por su habilidad en hacer la ciencia accesible al público profano. Demostró Tyndall de qué modo un *smog* producido artificialmente adquiría un "color que rivalizaba con el más puro cielo italiano", cuando se iluminaba con un haz de luz blanca y se observaba desde un punto de mira inclinado con respecto al haz. A partir de entonces, y durante muchos años, fueron numerosos los investigadores, Rayleigh incluido, que creyeron que el azul del cielo lo producía la difusión de la luz por partículas. Más aún, llegaron a la conclusión de que un gas puro, el aire, limpio de partículas de todo tipo, sería incapaz de difundir la luz y descomponerla en sus colores.

Pero, por fin, en un trabajo publicado en 1899, Rayleigh afirmaba que la difusión y separación de colores se debían a las propias moléculas del aire y que, "incluso en ausencia de partículas extrañas, el cielo seguiría siendo azul". Para entonces, había ya elaborado un elegante modelo que explicaba cómo difundía la luz una molécula. Para entender este modelo, consideremos una molécula de aire (no importa de qué clase), iluminada por luz solar blanca. Se halla ésta compuesta por todos los colores del espectro visible, cada uno de los cuales posee una longitud de onda asociada. La longitud de onda aumenta desde el azul al verde, y de éste al amarillo y rojo, de tal manera que la longitud de onda asociada a la luz roja es unas 1,68 veces mayor que la longitud de onda asociada a la luz azul.

Una molécula difunde cada color componente de la luz solar en todas las direcciones, pero no con una intensidad uniforme. La difusión más fuerte tiene lugar hacia adelante (como si la luz atravesara la molécula) y hacia atrás (en dirección al sol). La luz difundida en ángulo recto con el trayecto inicial de la luz solar es sólo de intensidad mitad. Todos los colores se difunden de acuerdo con este modelo, pero la intensidad que se difunde en cada dirección particular es función del color. Rayleigh descubrió que la intensidad dependía de la inversa de la cuarta potencia de la longitud de onda. Por tanto, una luz de longitud de onda corta (el azul, por ejemplo) se difunde con intensidad mayor que una luz roja, pongamos por caso, pues la longitud de onda de ésta es más larga. Por ser de 1,68 el cociente entre sus longitudes de onda, la luz difusa azul será de $1{,}68^4$ ve-

ces (ocho, aproximadamente) más intensa que la luz difusa roja.

Supóngase que interceptamos la luz que se difunde lateralmente formando un ángulo de unos 90° con la dirección inicial de la luz solar. Si pudiéramos percibir la luz procedente de una molécula aislada, la veríamos azulada, puesto que la mayor intensidad la produce el extremo azul del espectro. La situación adquiere mayor veracidad cuando toda una multitud de moléculas difunde la luz hacia nosotros, y se nos deja ver la luz y su color. Tal es el caso cuando levantamos la vista hacia un segmento de cielo alejado del sol. Entonces, todas las moléculas situadas en nuestra visual difunden hacia nosotros una luz cuya intensidad está dominada por el azul, y ese segmento de firmamento parece azulado; aunque no de un azul puro, ya que interceptamos también los otros colores, si bien éstos son más tenues.

El hecho de que el extremo azul del espectro sufra una fuerte difusión hacia el exterior del haz luminoso inicial implica que, en el haz que sigue camino por la atmósfera, va predominando gradualmente el extremo rojo del espectro. Si miramos al sol cuando éste se encuentra alto, la luz que nos llega ha viajado demasiado poco por la atmósfera para verla apreciablemente enrojecida. Cuando el sol está bajo y la luz solar recorre un trayecto más largo por la atmósfera, hasta llegar a nosotros, se muestra sensiblemente enrojecida; por ello, predomina el extremo rojo del espectro en los ocasos.

Alguien podría denunciar, en este razonamiento, una contradicción evidente. La luz azul sufre siempre una difusión más intensa que la roja, en cualquier dirección. Esto es cierto incluso para la luz que se difunde hacia adelante y continúa en la dirección inicial del haz. Así, pues, si la luz azul se difunde hacia adelante más intensamente que la roja, ¿por qué se enrojece progresivamente el haz?

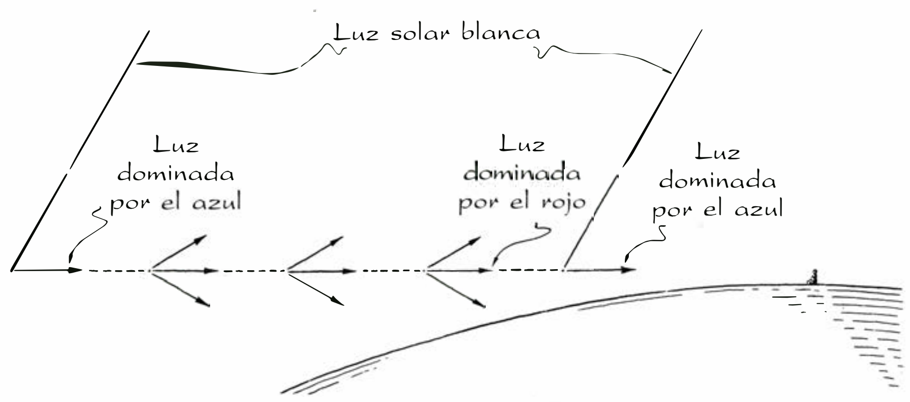
James A. Lock, colega mío en la Universidad de Cleveland, recurre al modelo corpuscular de la luz para mostrar que no hay tal contradicción. Supongamos que el haz de luz inicial posea 1000 fotones azules y otros 1000 rojos. Cuando el haz alcanza un grupo de moléculas, el número de fotones azules que se difunden en todas las direcciones (o en cualquier dirección concreta) octuplica el número de fotones rojos. Supongamos que en todas las direcciones se difundan 80 fotones azules y 10 fotones rojos, de los cuales se difunden hacia adelante 8 fotones azules y un fotón rojo. Entonces, al haz que sigue su camino, le quedan 991 fotones rojos, pero sólo 928 fotones azules, y por ello se enrojece.

Podríamos preguntarnos asimismo de qué modo puede difundirse la luz hacia adelante, atravesando una molécula. Lo que sucede es que ésta no constituye una barrera maciza a la manera de un muro, sino que presenta en su mayoría un espacio desocupado, con electrones situados en diversas órbitas, en torno a un núcleo considerablemente reducido. Según el modelo clásico de la difusión, el campo eléctrico de la luz insta, a su paso, la oscilación de los electrones, oscilaciones que se desarrollan a costa de energía que cede el campo eléctrico. Cuando se obliga a oscilar a una partícula del estilo del electrón, emite luz en todas las direcciones, salvo en la dirección de oscilación. Esta luz emitida es la luz "difundida" y, por supuesto, parte de ella sale hacia adelante.

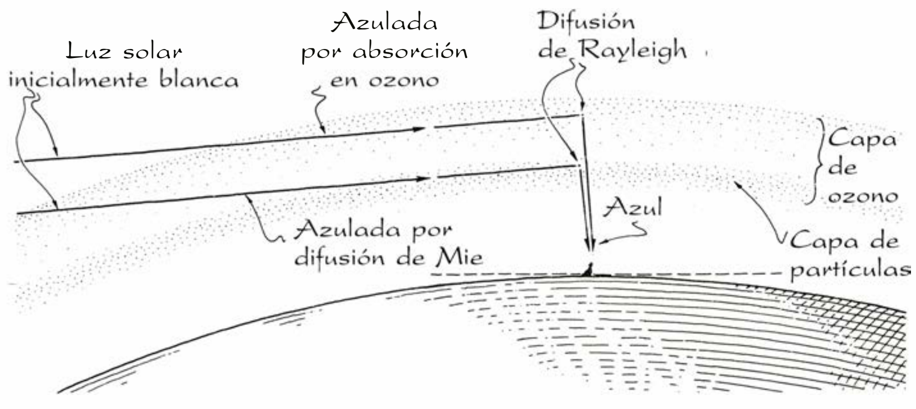
Bohren y Fraser presentaron cierta objeción a la explicación de Rayleigh sobre el azul del cielo, que luego retiraron. De todas las longitudes de onda que componen el espectro visible, las más cortas corresponden al violeta, no al azul. ¿Por qué, entonces, no es violeta el cielo? Bohren y Fraser aportan dos razones. Una, de secundario interés, es la siguiente: puesto que la luz solar inicial tiene un contenido de violeta algo más débil que de azul, se difunde, hacia nosotros, menos violeta que azul. Más importante es esta otra causa: el ojo humano es mucho menos sensible al violeta que al azul.

La gente atribuye a veces el azul del cielo al vapor de agua de la atmósfera, acaso porque las masas de agua suelen ser azuladas. Ahora bien, el motivo de que un lago nos parezca azul estriba en que, cuando una luz blanca atraviesa varios metros de agua, las moléculas de ésta absorben parcialmente el extremo rojo del espectro y la luz que al final se refleja hacia el observador resulta esencialmente azul. Pero Bohren y Fraser señalan que la atmósfera carece de agua suficiente para que ese tipo de absorción desempeñe un papel decisivo en el azulamiento del cielo.

El color del firmamento se ha atribuido también a la capa de moléculas de ozono que ocupa desde unos 10 hasta unos 40 kilómetros de altitud, y que alcanza su máxima densidad a unos 25 kilómetros. Estas moléculas poseen sus bandas de absorción en el extremo rojo



4. Por qué el cielo diurno es blanco junto al horizonte



5. Por qué se intensifica el azul en el cenit después del ocaso

del espectro. Podría ser que los componentes rojos de la luz solar se debiliten cuando ésta atraviesa la capa de ozono y que, por ello, predomine el azul en la luz que finalmente llega al suelo. Bohren y Fraser sostienen que, aceptado que el ozono empobrece el contenido de rojo, su papel causal en la coloración azul del cielo es secundario. Cuando miramos al cielo diurno, interceptamos una luz que ha atravesado muy poco ozono para que la absorción se note. En el crepúsculo, cuando los rayos nos llegan siguiendo un trayecto sesgado (y, por tanto, más largo) por el interior de la capa de ozono, la absorción en éste reviste mayor importancia, pero aún entonces la causa principal del azul del cielo hay que buscarla en el mecanismo de Rayleigh.

Durante el día, el azul del cielo empalidece en las cercanías del horizonte; en una franja de unos cinco grados a partir de éste, el cielo suele mostrarse blanco. Las moléculas de aire situadas a lo largo de una visual hasta el horizonte deben, como las otras, difundir la luz obedeciendo el mecanismo de Rayleigh; ¿qué puede haber ocurrido, pues, con el color azul? Para Bohren y Fraser, esa falta de color se debe al largo trayecto que en el seno de la atmósfera recorre la luz que llega a nosotros cuando miramos al horizonte, y no verticalmente. Con esa distancia adicional, la luz se difunde muchas más veces antes de llegar al lugar donde nos hallamos.

Parte de la luz se difunde en moléculas no muy lejanas [*véase la figura 4*]. De éstas recibimos una luz en la que predomina el azul. Las moléculas más distantes difunden también hacia nosotros una luz en la que predomina el azul; mas, por la lejanía, esa luz sufre repetidas difusiones antes de llegarnos. En cada proceso de difusión, la luz emitida hacia nosotros es luz difundida en dirección hacia adelante, por lo que se empobrece en azul; tras muchos procesos de difusión, esa luz acaba dominada por la mitad roja del espectro. En virtud de ello, de las moléculas más próximas recibimos luz comprendida principalmente en la mitad azul del espectro y, de las moléculas más lejanas, recibimos luz comprendida principalmente en la mitad roja. La combinación de ambas da blanco, que es el color que apreciamos en la dirección del horizonte.

El mismo efecto explica el color de las montañas en los días despejados. Si las montañas no quedan muy lejos, su imagen será azulada, porque las moléculas del aire situado entre aquéllas y nosotros difunde una luz en la que predomina el azul. Unas montañas algo más lejanas podrían parecernos más azules incluso; pero si esa lejanía crece, nos parecerán blancas, lo mismo que el horizonte.

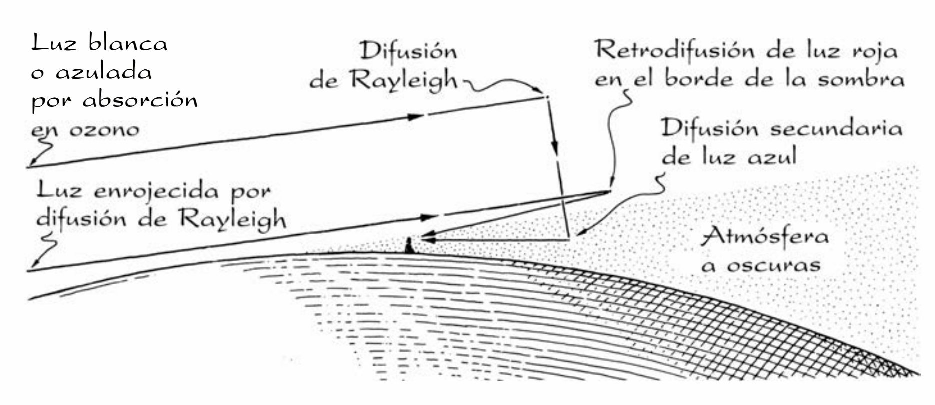
Según Bohren y Fraser, la luz procedente de un sol de poniente sería anaranjada (entre roja y amarilla), y no roja, si en su camino hacia nosotros se difundiera únicamente en las moléculas del aire. Ahora bien, señalan, la razón de que se nos ofrezca de un rojo vivo obedece a que no sólo se difunde en moléculas, sino también en minúsculas partículas y aerosoles contenidos en la atmósfera.

Cuando, a cualquier hora del día, levantamos la mirada a una región cercana al sol, interceptamos parte de la intensa luz que emite, difundida ya hacia adelante por esas partículas y aerosoles; por ello, esa zona se nos presenta más luminosa que si no hubiera tales partículas. Cuando el sol está alto, su entorno se muestra de un blanco intenso; cuando está bajo, por contra, la luz que llega a las partículas ya se ha enrojecido mediante la difusión de Rayleigh y su entorno se muestra rojo subido. Cuanto mayor sea la densidad de partículas, tanto más brillará el color del entorno y tanto más nítido aparecerá el perímetro del sol en el ocaso.

He venido suponiendo que el diámetro de las partículas y aerosoles no llegaba a 0,1 micras y que, en consecuencia, difundían la luz por el mismo mecanismo de Rayleigh que las moléculas. Ahora bien, las partículas de tamaño superior a 0,1 micras difunden la luz por un mecanismo mucho más complicado, llamado difusión de Mie, en honor de Gustav Mie, físico alemán que, a principios de siglo, elaboró un modelo teórico para este tipo de difusión. La difusión de Mie, producida por partículas de tamaño algo mayor, es realmente una forma de difracción en que la luz sale, en su mayor parte, hacia adelante, dentro de un cono estrecho. En este caso, la luz roja se esparce por un cono más ancho que la luz azul y, por tanto, el haz que sigue su camino se torna más azul. En la difracción simple, las ondas luminosas que inciden sobre una partícula se esparcen y, además, se propagan rodeando la partícula y penetrando en la "zona de sombra" de la misma. No entraremos en los detalles de la difusión de Mie, de interpretación más compleja.

Durante el crepúsculo de un atardecer despejado, el cenit (el punto del cielo situado exactamente en nuestra vertical) se vuelve más azul que durante el día. Ese azul adicional parece extraño al considerar que el horizonte puede estar muy rojo en la proximidad del sol. Se han propuesto varias explicaciones para este fenómeno, y en la más plausible de ellas interviene la capa de ozono. Cuando la luz procedente del sol, durante el ocaso, recorre un trayecto sesgado a través de la capa de ozono, la absorción que hace éste del extremo rojo del espectro deja al haz dominado por el extremo azul, pese a la difusión de Rayleigh que la luz sufre en su camino. En el haz predomina todavía más el azul si el mismo, además, recorre sesgadamente una capa de partículas de ozono lo bastante gruesa para introducir la difusión de Mie. Después de que la luz se haya vuelto azulada por uno u otro mecanismo, o por ambos, parte de ella se difunde hacia nosotros desde el cenit, mediante el mecanismo de Rayleigh; vemos así llegar, desde ese lugar, un azul más puro que el diurno.

Tras el ocaso, la sombra de la tierra comienza a alzarse por levante. El borde de esa sombra acostumbra a ser rojo o púrpura rosado. Este color se debe a la luz enrojecida por difusión de Rayleigh durante su larga travesía por los



6. Colores que se ven en el borde superior de la sombra de la tierra

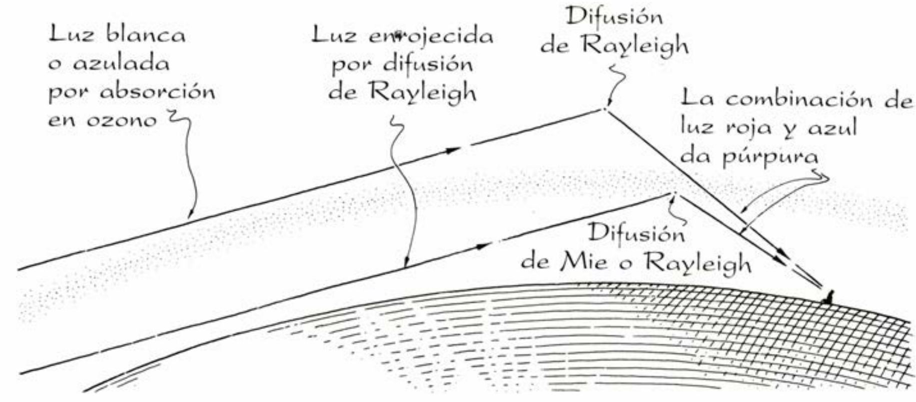
tramos de la atmósfera de altitud más baja. Allí donde distinguimos el "borde" superior de la sombra, parte de la luz sufre una difusión de Rayleigh y retorna hacia nosotros. Cuando interceptamos esa luz, percibimos rojo el borde superior.

Por debajo del borde, la parte superior de la sombra pudiera adoptar un tono azul pálido. Este tinte azulado se debe probablemente a la luz solar que atraviesa zonas de la atmósfera más altas y menos densas, donde la componente azul del haz no se debilita tanto como en un haz que pase por zonas menos altas de la atmósfera, más densas en moléculas de aire. En realidad, la luz podría azularse si atravesara sesgadamente la capa de ozono o una capa de partículas que forzara la difusión de Mie. En las proximidades de la sombra de la tierra, parte de la luz sufre una difusión de Rayleigh y penetra en la sombra; sufre allí otra difusión de Rayleigh, antes de encaminarse hacia nosotros. Como en este proceso participan múltiples difusiones, la luz procedente de la sombra que interceptamos es pálida, y, si la percibimos, es porque contrasta con la sombra oscura de la tierra.

A veces, unos diez minutos después de la puesta del sol, se genera sobre éste una mancha púrpura, entre 35 y 75 grados desde el cenit. Esa "luz púrpura", como suele llamarse, parece depender de la presencia de una capa de partículas a una altitud de 16 a 20 kilómetros, en la parte inferior de la capa de ozono. Tales partículas pueden ser polvo de algún desierto o ceniza muy fina procedente de erupciones volcánicas o grandes incendios forestales.

En 1967, Aden B. Meinel y Marjorie Pettit Meinel, de la Universidad de Arizona, hicieron notar que la capa pudiera estar formada por algún producto de erupciones volcánicas. Si una erupción emite una gran cantidad de dióxido de azufre, cuando el gas se mezcla con la base de la capa de ozono, reacciona con éste produciendo sulfatos. Y cuando éstos precipitan formando núcleos de condensación, crean un aerosol que actúa de difusor de Mie.

En *Sunsets, Twilights, and Evening Skies*, un libro delicioso, los Meinel sugieren que la mancha púrpura es el resultado de una luz muy roja y una luz muy azul que se difunden desde varias zonas del firmamento. La componente roja procede de luz solar que bordea la tierra, atravesando tal porción de atmósfera que la difusión de Rayleigh la torna roja. Parte de esta luz se difunde hacia nosotros desde la capa de partículas, cabe pensar que por el mecanismo de Mie, si las partículas son lo bastante grandes, y por el mecanismo de Rayleigh, si las partículas son más pequeñas; cualquiera que sea el caso, recibimos luz roja suplementaria por mor de las partículas.



7. Formación de la mancha púrpura en el crepúsculo

La componente azul procede de la luz solar que atraviesa unas zonas más altas de la atmósfera y, por ello, no se enrojece tanto. (Podría añadirse que esa luz atraviesa sesgadamente la capa de ozono; podría, pues, estar dominada por el extremo azul del espectro a causa de la absorción.) Parte de esa luz sufre difusión de Rayleigh y recibimos luz azul. Entonces, cuando miramos hacia la mancha, las dos componentes, la roja y la azul, viajan a lo largo de nuestra visual, y su combinación crea la impresión de luz púrpura.

La causa de que otras porciones del firmamento no parezcan púrpuras estriba en que nos envían diferentes mezclas de colores, y no sólo rojo y azul, cuando miramos hacia ellas; pueden ofrecernos toda una variedad de tonalidades en función de nuestro ángulo de visión. Los colores resultan especialmente brillantes cuando la capa de partículas es densa y extensa, como suele ocurrir tras una erupción volcánica de importancia. Así, en 1883, la explosión del Krakatoa, en Java, produjo unos ocasos brillantes durante unos cinco años, y la del Agun, en Bali, ocurrida en 1963, repitió el espectáculo a lo largo de tres años.

Seguimos sin conocer la razón de una segunda luz púrpura (aunque mucho más infrecuente): aparece aproximadamente en la misma posición del firmamento que la primera, entre hora y media y dos horas tras el ocaso. Los amantes de los crepúsculos opinan que se debe a la misma capa de partículas que la primera luz púrpura. En su opinión, si la capa es extensa, parte de la luz que se difunda desde la porción de la capa situada muy por debajo del horizonte puede volver a difundirse desde la porción de la capa que esté a la vista. Entonces, si la luz brilla, veremos una mancha de color púrpura pálido.

Otra explicación recurre a una segunda capa de partículas situada a una altitud de 80 a 90 kilómetros, en una región de baja temperatura ubicada donde se unen la mesosfera y la ionosfera (ésta se encuentra inmediatamente encima de la anterior). Las partículas podrían ser de origen terrestre, pero es más probable que provengan del espacio; la tierra intercepta una enorme cantidad de cometas y restos de asteroides que podrían producir a veces una extensa capa de partículas en la región de baja temperatura. Desde luego, la luz solar difundida por tal capa es demasiado apagada para que la percibamos durante el día, o incluso al comienzo del crepúsculo; podría, en cambio, dejarse ver con una atmósfera que se encuentre en la sombra de la tierra y la capa siga iluminada por el sol. La creciente contaminación luminosa debida a los ambientes urbanos podría perfectamente hacer aún más infrecuente en el futuro la observación de esta segunda mancha púrpura.

BIBLIOGRAFIA COMPLEMENTARIA

POLARIZATION AND SCATTERING CHARACTERISTICS IN THE ATMOSPHERES OF EARTH, VENUS, AND JUPITER. David L. Coffeen en *Journal of the Optical Society of America*, vol. 69, n.º 8, págs. 1051-1064; agosto, 1979.

RAYLEIGH SCATTERING. Andrew T. Young en *Physics Today*, vol. 35, n.º 1, págs. 42-48; enero, 1982.

SUNSETS, TWILIGHTS, AND EVENING SKIES. Aden Meinel y Marjorie Meinel. Cambridge University Press, 1983.

COLORS OF THE SKY. Craig F. Bohren y Alistair B. Fraser en *The Physics Teacher*, vol. 23, n.º 5, págs. 267-272; mayo, 1985.

Juegos de ordenador

Problemas gentilicios: tema y variaciones

A. K. Dewdney

"El análisis del carácter parece ser el más excelso entretenimiento humano."
—Isaac Bashevis Singer, *New York Times Magazine*, 26 de noviembre de 1978

Tres filósofos de la antigua Grecia fueron después de comer a dar un paseo campestre por los alrededores de Atenas. Hallaron sombra al pie de un olivo, se acomodaron allí, descorcharon una ánfora de vino y se pusieron plácidamente a analizar la cuestión ontológica fundamental, a saber: ¿por qué existe algo? La discusión se acaloró, se tornó en confusión y acabó en disputa. Al poco, los tres filósofos se quedarón dormidos a la sombra del aceituno.

Avanzada la tarde, un jovencito ateniense de no muy santas intenciones estuvo espiando a los tres dormidos filósofos. Con gran cuidado, salpicó sus frentes con unas gotas de pintura blanca. Justo antes de la puesta de sol, un búho que había hecho del árbol su domicilio se posó sobre una rama, encima mismo de los tres hombres. El búho ululó sonoramente y se alejó volando con gran batir de alas. El grito del búho hizo despertar a los filósofos, cada uno de los cuales decretó inmediatamente que el búho era responsable de los ornamentos que sus dos colegas exhibían en sus frentes. Todos soltaron una carcajada.

No cabía duda de que se trataba de una visión muy divertida, pues pasaron unos cinco segundos antes de que uno de los tres interrumpiera bruscamente su reír. ¿Por qué? Podemos resolver el problema reconstruyendo el proceso mental del filósofo que cesó en sus risas. Al hacerlo, uno se ve llevado a considerar qué debió suponer él que estaban pensando sus colegas.

Son muchos los rompecabezas donde intervienen personas, sin más excusa que proporcionar a sus hipotéticos resolvientes un contexto en el que sentirse más cómodos. Pero las soluciones de los problemas que llamo "rompecabezas gentilicios", como el de los tres filósofos, se fundan en la reconstrucción de lo que cada uno de quienes en ellos intervienen piensa que están pensando los demás. Tal razonamiento sobre el razonamiento no queda sólo en tema divertido para rompecabezas y charadas; les es también necesario a los científicos informáticos que se esfuerzan en crear programas capaces de remedar la forma en que un ser humano podría pensar en situaciones similares e igualmente desconcertantes.

El tema principal de este artículo no quedaría completo sin la solución del problema de los tres filósofos. Sean éstos Pitágoras, Platón y Aristóteles. Fue Pitágoras, que era de los tres el más viejo y sabio, quien interrumpió sus risas. He aquí la razón: al ver Pitágoras cómo se desternillaba Aristóteles, se percató de que éste no tenía ni idea de que su propia frente estaba untada con una sustancia blanca. Si su propia frente (la de Pitágoras) estuviese limpia, sería evidente que Aristóteles estaría riéndose de Platón. Pero entonces, ¿de quién pensaba Aristóteles que se reía Platón? "¡Por el casco de Atenea!", exclamó sin duda Pitágoras para sus adentros, "¡No debería estar riéndome!". La situación se expone gráficamente en la viñeta de la figura 1.

La explicación anterior es la solución de la versión más o menos tradicional del rompecabezas. El problema se puede generalizar un poco preguntando por qué dejó de reír Platón algunos segundos más tarde. En cuanto Pitágoras interrumpió sus risas, la serie de pensamientos de éste ya no le es evidente a Platón. En realidad, podría parecer que en la mente de Platón queda ahora reforzada la convicción de tener la frente limpia. "Es evidente que Pitágoras ha visto que tengo la frente limpia, y cree que Aristóteles está riéndose de él", podría pensar Platón. Mas si Platón llevase un poco más allá sus reflexiones, podría deducir su propio y embadurnado estado. Dejaré que el lector se calce las sandalias de Platón, para que determine cómo.

Fijémonos en una variante del tema de los tres filósofos. Un sultán quería elegir al más sabio de tres de sus visires, para nombrarle gran visir. El sultán llevó a los visires a una habitación oscura, y les puso a cada uno en la cabeza un fez blanco. Seguidamente les condujo a la sala del trono y les dijo: "Cada uno de vosotros lleva puesto un fez que puede ser blanco o negro; al menos uno de los que lleváis es blanco. Quien primero me diga el color de su fez será nombrado gran visir".

El problema de los tres visires es esencialmente el mismo que el de los tres filósofos, si bien ahora no hay risas que sirvan de indicador de lo que cada visir está viendo. De una forma curiosa, el conocimiento compartido de que al menos uno de los gorros es blanco coopera al alirón con el silencio de la indecisión, y produce efectos equivalentes.

Llamemos a los candidatos al-Khwarizmi, ibn Khaldun e ibn Sina. Estos fueron los pensamientos de al-Khwarizmi, que obtuvo el cargo:

"Veamos. Me pregunto si mi fez es negro. Si lo fuera, ¿qué podrían pensar los otros visires? Supongamos que ibn Khaldun estimara también que su gorro es negro. En tal caso se daría cuenta de que ibn Sina estaría viendo dos gorros negros, y deduciría inmediatamente que el suyo es blanco. Sin embargo, ibn Sina no ha exclamado '¡Mi gorro es blanco!'; por tanto, ibn Khaldun sabe que ibn Sina está viendo al menos un gorro blanco. Pero si realmente mi propio fez fuese negro, ibn Khaldun sabría que el suyo es el gorro blanco que ibn Sina está viendo, y así lo diría. Sin embargo, ibn Khaldun sigue callado. ¡Por las barbas del Profeta, mi gorro no puede ser negro!".

El problema de los visires tiene un nivel de metaproblema que me fue señalado por uno de mis colegas. Hélo aquí: en cuanto se les informa a los visires de la naturaleza de la prueba, cada uno de ellos puede deducir que la única prueba equitativa de sus facultades cognitivas requiere forzosamente que todos los gorros sean blancos.

No hay dificultad en reformular el problema de los tres filósofos o los tres visires de un sinfín de formas. En una variación bien conocida, tres aristocráticas damas inglesas de fines del siglo pasado viajan en ferrocarril. El día es caluroso y llevan abierta la ventanilla de su departamento, para dejar entrar un poco de aire fresco. La locomotora, eruptando espeso humo negro, penetra en un largo túnel. Cuando el tren sale de nuevo a la luz del día, las tres mujeres se echan a reír, al ver los rostros cubiertos de hollín de sus compañeras de viaje... hasta que una de ellas deja de hacerlo.

He aquí una segunda variante: tres hombres de negocios están almorzando un pastel de espinacas en un restaurante de moda neoyorkino. A todos les quedan hebras de espinaca entre los dientes... La comprensión de que un rompecabezas "gentilicio" puede adoptar multitud de formas equivalentes conduce a la idea de la transformación de rompecabezas por "isomorfismo". ¿Cómo transformar, por así decirlo, filósofos en visires? Resulta conveniente identificar los elementos de cada problema que tienen en el otro elementos homólogos:

1. joven bromista → sultán
2. filósofo → visir
3. frente → cabeza
4. manchón blanco → fez blanco
5. reír → callar

Además de estos elementos, la transformación debe definir también la característica peculiar de un filósofo o visir tal cual está operacionalmente definido en el problema. Un filósofo es alguien que, hasta que analiza la situación, seguirá riéndose si ve al menos una frente pringada, y no se reirá en caso contrario. Un visir es alguien que, hasta que analice la situación, no dirá nada si ve al menos un gorro blanco. En caso contrario dirá: "Llevo puesto en la cabeza un fez blanco".

Invitamos a los lectores a rellenar los espacios en blanco del rompecabezas gentilicio genérico que damos a continuación. Al rellenar los espacios numerados con las correspondientes palabras de la columna de la izquierda de la tabla anterior, se obtiene el problema de los filósofos. Haciéndolo con las palabras correspondientes de la columna derecha, hace aparición el de los tres visires. (Evidentemente, uno puede rellenar los huecos con una serie propia de palabras que vengan a cuento.) Para mayor simplicidad, se han eliminado los elementos pintorescos de tiempo y lugar.

En cierta ocasión, un (1)________ puso un (4)________ en la (3)________ de tres (2)________ (s) sin que éstos lo supieran. Mientras una de estas personas no supiera que su propia (3)________ tenía un (4)________, cada (2)___ no dejaría de (5)___. De repente, el más sabio de los tres (2)________s, deja de (5)________. Entonces exclama: ________. "¡Mi (3) ________ tiene un (4) ________!". ¿Cómo pudo deducir tal cosa?

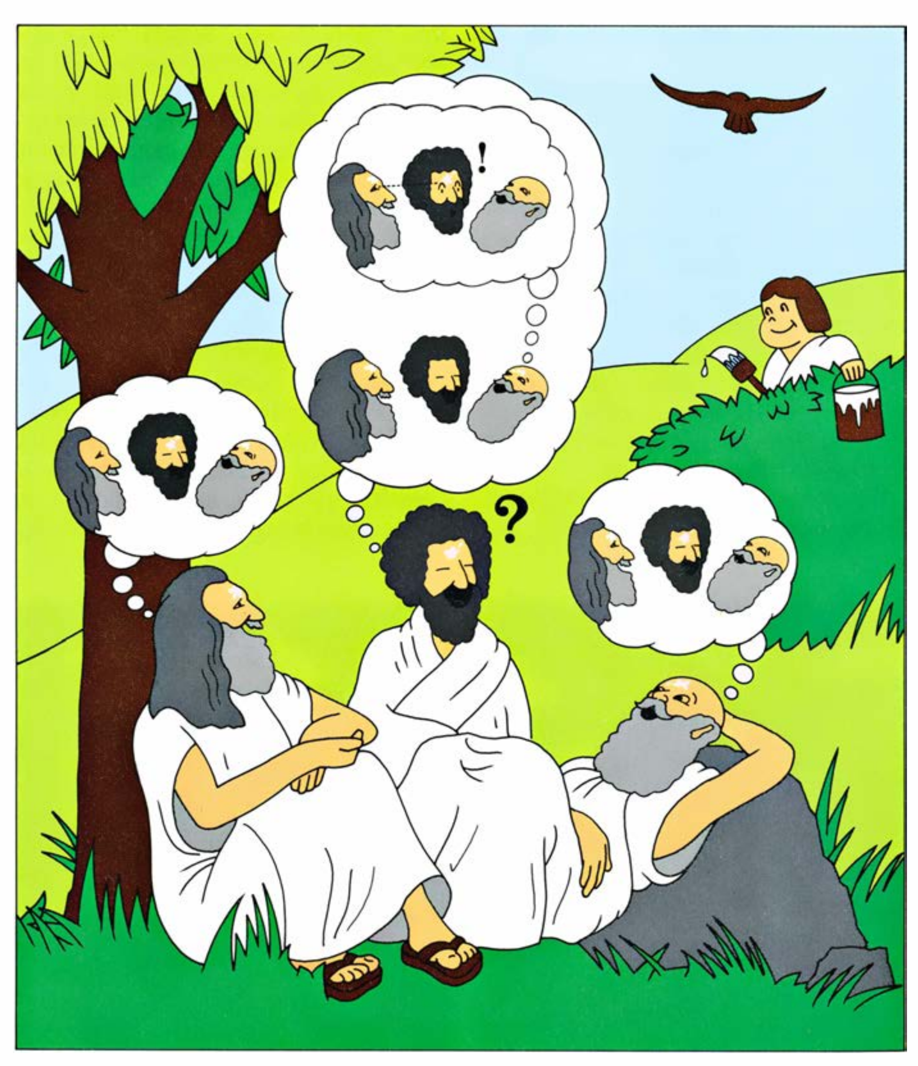
En la versión original de la variación siguiente sobre el tema de los tres filósofos, los personajes eran esposas infieles. Para evitar estereotipos sexistas, la someteré a una sencilla transformación. En cualquier caso, el problema es interesante porque generaliza para un número cualquiera de individuos el problema de los tres filósofos.

La tiránica reina de las Amazonas anunció un día a sus súbditas que al menos uno de los hombres de su reino había sido infiel. Lanzó entonces un edicto: "Ordeno a cualquiera de vosotras que descubra que su marido le es infiel que, al sonar las campanadas de medianoche del día en que esté segura de su infidelidad, le ejecute sin piedad". En el reino de las Amazonas la información se compartía libremente, aunque no con excesiva libertad: cada esposa conocía las infidelidades de todos los demás maridos, pero no las del suyo propio. Por otra parte, la noticia de que un marido había sido ejecutado se difundía por todo el reino antes de concluido un día. Resultó que los maridos infieles eran 40. ¿Fue ejecutado alguno de ellos?, y de ser así, ¿cuándo?

Los lectores se habrán fijado en que la reina anunció que al menos uno de los maridos había sido infiel. Si exactamente un marido hubiera sido infiel, la mujer lo habría sabido inmediatamente: si alguno de los maridos, salvo el suyo, hubiera ido de picos pardos, ella lo habría sabido ya. Por consiguiente, llegada la medianoche del día del edicto real, el pícaro hubiera sido ejecutado por su esposa.

Si fueran exactamente dos los esposos infieles, sus respectivas esposas les hubieran despachado a la medianoche del segundo día. La noticia de que la primera noche no se había producido ejecución alguna les habría confirmado que dos de los maridos eran infieles. Dado que las esposas de los dos tenorios solamente habrían tenido noticia de uno de ellos (aunque todas las demás amazonas sabrían que había al menos dos), aquéllas se habrían dado cuenta de que su propio esposo tendría que ser el segundo de los cupidos locales.

A estas alturas, los lectores ya se habrán dado cuenta de la línea del razonamiento. La ausencia de ejecuciones



1. Uno de los filósofos está razonando en un plano más profundo que sus colegas

en la medianoche del n-ésimo día es signo de que al menos n+1 de los esposos habían sido infieles. Al amanecer del cuadragésimo día, sería de conocimiento común que al menos 40 maridos eran infieles. La verdad es que ello no sería sorpresa para las amazonas casadas con maridos fieles, pues éstas sabían ya que eran 40 los que sacaban los pies del tiesto. Solamente las esposas de los adulterinos sabrían de 39, lo que significaba que el suyo propio era el cuadragésimo. Tales esposas, al cumplir la orden de su monarca, convocarían a sus cónyuges para un tête a tête de medianoche al cumplirse los cuarenta días del edicto real.

¿Es verdaderamente el rompecabezas de las Amazonas una variante del problema de los tres filósofos? Podemos responder preguntándonos qué sucedería si solamente hubiera tres amazonas, casadas cada una de ellas con un esposo infiel. En ese caso, al final del tercer día, cada amazona habría deducido la conclusión correcta. Así planteado, el problema se asimila muy directamente con el de los tres visires (que, como ya he hecho ver, es una transformación del problema de los tres filósofos).

Para ver que así es, supongamos que el sultán hubiera dicho a los candidatos a gran visir: "Os preguntaré varias veces en sucesión si conocéis el color del fez que lleváis en la cabeza. Responded solamente si estáis seguros; de lo contrario, guardad silencio". En este caso, la primera vez que el sultán hiciera la pregunta, los tres visires hubieran callado. Al preguntarlo por segunda vez, los tres hubieran vuelto a guardar silencio. Y al ser interrogados por tercera vez, los tres candidatos hubieran respondido afirmativamente.

En el caso de las amazonas se trata de 40 esposos infieles, no de tres. ¿Podríamos generalizar el problema de los tres filósofos a un grupo de 40? Sí, en efecto. Supongamos, por el momento, que sean cuatro los filósofos dormidos bajo el árbol. Al despertar todos se echan a reír, y el cuarto filósofo (en realidad, un dios que ha tomado apariencia humana) razona como sigue:

"Vamos a ver. Es acorde con mi divina dignidad suponer que llevo la frente inmaculada, por lo que puedo mofarme de estos tres mancillados mortales. Mas, ¿por qué no se percata ninguno de ellos de mi condición y deja de reírse?". (El olímpico visitante está recapitulando ahora el razonamiento de Pitágoras.) "¡Ay de mí! ¡Me parece saber el motivo!"

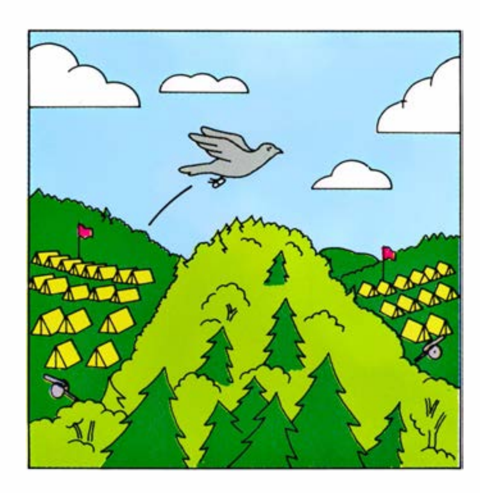
Si semejante proceso mental puede llevar a un cuarto filósofo a la comprensión de que tampoco su frente ha sido respetada, entonces servirá igualmente bien para explicar por qué un quinto, un sexto e incluso un cuadragésimo filósofo podría llegar a idéntica conclusión. En una de sus secciones de "Juegos matemáticos", Martin Gardner desarrolló en esta misma línea una generalización del rompecabezas de los tres visires [véase "Juegos matemáticos", INVESTIGACIÓN Y CIENCIA, julio de 1977]. Sin embargo, como sagazmente hacía notar Gardner, surgen dificultades. "Normalmente esta generalización es puesta en tela de juicio. El problema exige tantas y tan nebulosas hipótesis sobre grados de inteligencia, y lapsos de tiempo razonables cada vez mayores, que acaba siendo irreal."

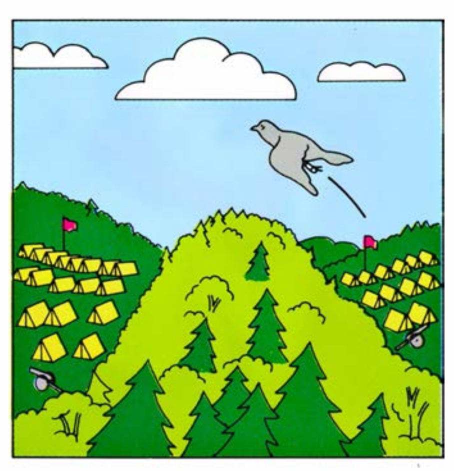
No hay duda de que todos los rompecabezas comportan un cierto grado de irrealidad. ¿Cuán probable es, incluso en un restaurante de buen tono, que tres personas que almuercen pastel de espinacas y estén sorbiendo un Chardonnay fresco, acaben todas a la vez con una pizca de espinaca entre los dientes? Y aun cuando así suceda, ¿qué probabilidad hay de que uno de ellos no se percate por la mirada de otro que también se estaban riendo de él?

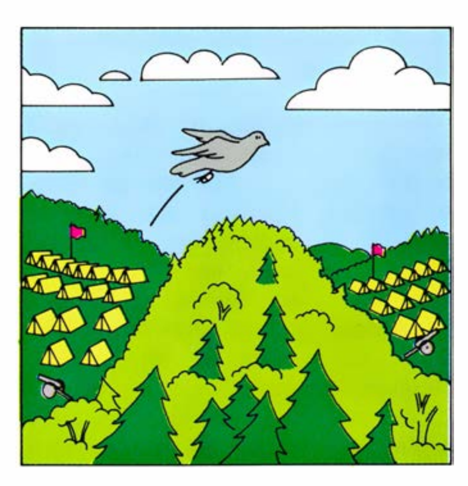
Sin embargo, para los investigadores que se esfuerzan por desarrollar nuevas formas de inteligencia artificial basadas en lo que se conoce como programación lógica, estos rompecabezas gentilicios son asuntos serios. John L. McCarthy, de la Universidad de California en Berkeley, entre otros, puso a prueba la potencia deductiva de los sistemas lógicos aplicándolos a problemas de personas como el de los tres visires. Los sistemas de programación lógica se valen de una categoría de procesos de pensamiento simbólico conocidos por cálculo de predicados, con el propósito de inferir automáticamente diversas deducciones a partir de ciertas hipótesis. Al objeto de remedar las facultades deductivas de los humanos en diversas situaciones sociales, los sistemas de programación lógica han de poseer potencia suficiente para modelizar razonamientos referentes a razonamientos.

Tales investigadores obrarían prudentemente examinando el tipo mucho más verosímil de problemas surgidos entre personas que ofrece la vida cotidiana. Aunque tales problemas no sean de gran profundidad lógica, y a pesar de que su resolución requiera multitud de hipótesis, revelan, empero, que una importante fracción de nuestra capacidad mental se aplica a reflexionar sobre lo que otras personas puedan estar pensando que pensamos nosotros.

Erwin Goffman, también de Berkeley, es uno de los pioneros en una especialidad de la psicología conocida por análisis transaccional, la cual sostiene que cada persona está tratando constantemente de saber lo que otras piensan de ella y manipular tal opinión, de forma muy similar a como hacen los







2. ¿Llegarán a tener fin las confirmaciones y reconfirmaciones entre los dos generales?

actores. Goffman ha descubierto este aspecto de nuestras vidas interiores por compilación de una serie de conductas "demasiado humanas" que todos reconocemos como experiencias válidas.

He aquí dos ejemplos del tipo descrito por Goffman, procedentes de mi propia vida, pues se da la circunstancia de que yo he estado recopilando ejemplos similares durante muchos años. Todos parecen apuntar a una capacidad de los humanos para pensar en lo que otras personas están pensando, facultad que en casi todos nosotros, estoy convencido, es semiconsciente.

En cierta ocasión, vi a un policía bajar hasta la calle desde un establecimiento de cuestionable naturaleza instalado en un segundo piso. Posiblemente se tratase de un local de masajes; la zona de la ciudad donde ocurrió el incidente lo hace verosímil. Al salir a la calle, el policía se tropezó inesperadamente con un compañero. Al ver a su colega, la cara que puso fue de lo más rara: su aspecto no era el de estar avergonzado, sino como si estuviera fingiendo estar avergonzado. Encontré al principio extraña tal conducta, pero luego me pareció más comprensible. El sentimiento primario del primer policía seguramente fuese de vergüenza pura y simple. Después de todo, su colega sospecharía inmediatamente que su visita había sido, por así decirlo, de carácter extraoficial. Sin embargo, si su rostro hubiera expresado la vergüenza que sentía, habría confirmado las sospechas del otro. Por otra parte, si el primer agente pudiera hacer que el segundo percibiera su vergüenza como fingida, es posible que no le diera al incidente más importancia que a una broma.

No pocas veces, los juegos hacen aflorar lo peor de las personas. Dos colegas francamente competitivos acababan de terminar una partida de póker. El perdedor estaba visiblemente molesto, aunque hacía violentos esfuerzos por contenerse. Finalmente reventó: "Tú estabas planeando ________me y yo estaba planeando ________te. ¡Sólo que has sido tú quien me ha llevado al huerto!" (Los huecos pueden sustituirse por palabras gruesas idóneas, que en este contexto significan hacer trampa.)

La palabra clave es sólo. A primera vista parece dar a entender que la victoria del oponente se debió sobre todo al factor suerte. Sin embargo, en inspección más atenta, parece indicar que el perdedor quiere que el ganador piense que él (el perdedor) opina que el aspecto más significativo del juego fue la igualdad de sus motivos y no la desigualdad que la derrota comporta.

Me pregunto si habrá lectores dispuestos a enviar ejemplos propios. Serían de particular interés aquellos de contenido real un poco más cercano a los rompecabezas gentilicios tradicionales.

Y hablando de rompecabezas, ha aparecido un maravilloso libro de problemas, titulado *The Puzzling Adventures of Dr. Ecco*. Su autor es Dennis Shasha, especialista en ciencias de cómputo del Instituto Courant de ciencias matemáticas de la Universidad de Nueva York. El libro, cuya traducción española sacará la editorial Labor en breve, describe las aventuras del misterioso Dr. Ecco, un excéntrico fabulosamente inteligente que se gana la vida resolviendo acertijos. Además de un par de problemas gentilicios, contiene rompecabezas electorales, rutas múltiples, espías, circuitos supervisores de otros y mucho más. Hay incluso un concurso donde se reta a los participantes a resolver 10 problemas de descodificación de textos. Los vencedores no sólo ganarán una "camiseta de omniheurista" (símbolo de suprema capacidad intelectual), sino también el derecho a que su nombre figure en un juego de piezas de ajedrez de ónice, grabado a mano. El concurso se cierra el 1 de abril de 1989. El ganador del concurso será anunciado en esta misma sección, en el artículo de septiembre.

Concluiré este comentario sobre problemas gentilicios con uno de los problemas del libro de Shasha. Dos generales, cuyos respectivos ejércitos ocupan laderas opuestas de una loma, quieren coordinar su ataque sobre un enemigo común, porque si cualquiera de sus ejércitos atacará solo, seguramente sería destruido. Desdichadamente, los generales nada más pueden comunicarse mediante palomas mensajeras, que han de ir de un campamento a otro sobrevolando la cima de la loma [*véase la figura 2*].

El primer general le envía al segundo el siguiente mensaje: "Ataque a las 8.00 horas. Confirme recepción mensaje; en caso contrario, no atacaré". El segundo general no tiene inconveniente en atacar a esa hora y, a punto ya de enviar una paloma con el acuse de recibo, se da cuenta de que el primer general no atacará a menos que reciba tal confirmación. Dado que el segundo general no tiene garantía de que la paloma mensajera llegue verdaderamente a entregar su mensaje al primer general, decide que no atacará hasta cerciorarse de que el primer general sabe que el ataque ha sido confirmado. El segundo general envía entonces otra paloma con un mensaje dirigido al primero. ¿Llegará a tener fin esta recurrencia infinita de mensajes confirmatorios? La respuesta depende tal vez del tipo concreto de mensaje que uno de los generales envíe en un determinado punto.

Reconozco y agradezco la ayuda que en la preparación de este artículo me han prestado dos de mis colegas de la Universidad de Western Ontario, Andy L. Szilard y Areski Nait Abdallah.

Anunciaré ahora quiénes fueron los vencedores del reto de los cuadros de primos, lanzado por Gordon Lee en el artículo de septiembre. Un cuadro de primos es un casillero de tamaño 6 × 6, cada una de cuyas cuadrículas contiene un número de una sola cifra. El objetivo consiste en elegir los dígitos y sus posiciones en el casillero, de modo que se obtenga el máximo número posible de números primos leyendo dígitos consecutivos a lo largo de cualquiera de las líneas rectas del casillero, sea en horizontal, en vertical o en diagonal. Lee, que organizó en Inglaterra un concurso similar el año pasado, informa que uno de tales cuadros —el del vencedor— contenía 170 números primos. ¿Podrían mejorar ese resultado los lectores de esta sección? ¡No lo dudé ni un instante!

Recibí muchos más cuadros de lo que me esperaba. Y los recibí tanto de adictos al ordenador con acceso a ordenadores de alta tecnología, cuando no superordenadores, como de personas de más humildes medios, que trabajaron a mano, con lápiz y papel. Le remití a Gordon Lee las matrices recibidas, para que él otorgase entonces los premios. He aquí los resultados, en orden inverso:

Matricero más tenaz: Larry J. Padden, de Oklahoma City, Oklahoma, que produjo 147 cuadros de primos diferentes, que contenían entre 170 y 173 primos cada uno.

Matriceros más "seixys". David Mckenzie y Frank Endres, de Austin, Texas, que descubrieron una matriz que, si bien sólo contiene 106 primos, la totalidad de los 28 números de seis cifras que figuran en ella son primos.

Tercer oficial de matricería: James I. Waldby, de Robinson, Illinois, cuya puntuación fue de 185.

Segundos oficiales de matricería: Mckenzie y Endres comparten este honor con Stephen C. Root, de Westboro, Mass. Todos ellos lograron dar con matrices de 186 números primos.

Matricero en jefe: Root, nuevamente. A continuación se da su matriz ga-

nadora, que contiene un total de 188 primos:

3	1	7	3	3	3
9	9	5	6	3	9
1	1	8	1	4	2
1	3	6	3	7	3
3	4	9	1	9	9
3	7	9	3	7	9

Fueron muchos los lectores que identificaron el camino más corto para atravesar el cruel laberinto descrito en el artículo de noviembre. Tiene una longitud de 39 casillas (contando las casillas de entrada y salida). Existen muchos otros, más largos, que atraviesan también el laberinto, pero todos los conducentes de una forma u otra hasta la salida han de pasar por un punto específico del interior del lóbrego laberinto. Es en ese punto donde el feroz Minotauro acecha a los infortunados jóvenes y doncellas atenienses que fueron obligados a entrar en el dédalo. Aunque algunos lectores descubrieron la secreta ubicación del Minotauro mediante el dibujo laborioso de un mapa y la localización posterior del punto de embotellamiento, otros sospecharon que el alto grado de simetría que ofrecía el diseño del laberinto podría ser la clave para determinar el emplazamiento del Minotauro. Y así era.

El dédalo tridimensional tan sólo presenta asimetría en un lugar. Los lectores encontrarán, sin duda, de su agrado retornar al laberinto y descubrir que "falta" en él una pared entre las casillas (5,3) y (5,4) del segundo nivel. (Las coordenadas designan, respectivamente, el número de casillas situadas a la derecha del borde izquierdo del laberinto y el número de casillas en sentido descendente desde el borde superior.) ¿Por qué ha de estar el embotellamiento en el punto de asimetría?

BIBLIOGRAFIA COMPLEMENTARIA

THE PRESENTATION OF SELF IN EVERYDAY LIFE. Erving Goffman. Doubleday & Company, Inc., 1959.

THE PUZZLING ADVENTURES OF DR. ECCO. Dennis Shasha. W. H. Freeman and Company, 1988.

Libros

Paleoantropología, bioquímica, Plinio y repertorios de fuentes

Jordi Agustí, Claudi M. Cuchillo y Luis Alonso

PALEONTOLOGÍA HUMANA. Selección e introducciones de Emiliano Aguirre. Prensa Científica, S. A.; Barcelona, 1988.

Emiliano Aguirre es una figura de larga trayectoria en la paleontología española. El miembro más joven de la generación que cimentó la paleontología de vertebrados en este país (Crusafont, Villalta, Truyols), a él le corresponde el mérito, entre muchos otros, de haber creado y desarrollado la prolífica sección de paleontología del Museo Nacional de Ciencias Naturales. Aunque la contribución más conocida de Aguirre corresponde al estudio de los restos de Atapuerca, su inquietud por la paleontología humana viene de muy atrás. El lector no especialista puede recordar en este sentido la serie de capítulos redactados para el volumen colectivo *La Evolución* (B. A. C., 1968) o *El origen del hombre* (Salvat, 1973).

Hoy, al cabo de varios años, Aguirre nos ofrece de nuevo una atractiva aproximación al tema que, con el título de *Paleontología humana*, recoge una selección de artículos aparecidos entre 1978 y 1987 en la revista *Investigación y Ciencia*. La obra consta, además, de una introducción general y tres introducciones parciales a cargo del autor, para los tres grandes apartados en que se estructura el volumen. La temática de estas tres grandes divisiones podría esquematizarse como sigue: *Homo sapiens* (neandertales incluidos) y *Homo erectus*, por un lado, *Homo habilis* y *Australopithecus*, por otro, y, finalmente, hominoideos pre-*Australopithecus* (en una sección globalmente dedicada a la aparición del bipedismo). Esta estructura presenta una lógica a la vez cronológica y biogeográfica: *Homo erectus* y sus descendientes son formas cosmopolitas que caracterizan la fase de expansión euroasiática de nuestro género. Por otro lado, *Australopithecus* y *Homo habilis* pertenecen a una etapa exclusivamente africana de nuestra evolución. Finalmente, driopitécidos y sivapitecinos corresponden nuevamente a formas ampliamente extendidas por el Viejo Mundo. Curiosamente, el coordinador del volumen ha preferido presentarnos el panorama de la evolución humana en sentido inverso a la flecha del tiempo, es decir, empezando por *Homo sapiens* y culminando en las formas miocénicas y el subsecuente proceso de adquisición del bipedismo.

Entrando ya en materia, la obra se inicia con el artículo de E. Trinkaus y W. W. Howells sobre los neandertales. Esta variante temprana de *Homo sapiens*, de distribución perimediterránea, fue considerada durante largo tiempo como una rama abortiva de pesado esqueleto y torpes movimientos. Hoy día sabemos que esta visión debe ser corregida: los neandertales eran ciertamente formas robustas que, pese a sus caracteres plesiomórficos (o primitivos), presentaban numerosos caracteres apomórficos (o derivados). Su situación con respecto a lo que se conoce como "moderno *Homo sapiens*" sigue siendo confusa. Hasta hace unos años, las formas halladas en el Levante mediterráneo (Palestina) parecían testimoniar una relación directa entre ambas variantes. Por el contrario, Trinkaus y Howells se muestran contrarios a un origen "neandertaliano" del hombre moderno. Los hallazgos realizados en Sudáfrica y Africa oriental parecen confirmar su punto de vista (aunque Aguirre se muestra cauto en cuanto a la definitiva resolución de esta cuestión). En cualquier caso, la aparente coexistencia de morfotipos "modernos" y "neandertaloides" en Palestina sigue sin tener una explicación definitiva. Como siempre, la discutible y poco afinada cronología de la mayor parte de estos hallazgos plantea problemas a la hora de situarlos en una secuencia evolutiva. El trabajo de Trinkaus y Howells tiene un interés suplementario, por cuanto no sólo se centra en los aspectos puramente craneales de la evolución de este grupo, sino que, además, hace referencia al esqueleto postcraneal de los neandertales y a los aspectos biomecánicos implicados.

Por lo que hace a la cuestión de los llamados "pre-neandertales" (o formas predecesoras de *Homo sapiens* en Europa), Aguirre la aborda en la introducción a la sección y en su propio artículo sobre el "hombre de Atapuerca". Aguirre realiza, en este apartado, una detallada enumeración de los restos ibéricos del Pleistoceno medio e inferior. Una vez más, el aspecto más problemático, en este caso, atañe a la cronología relativa de los hallazgos. Así, el húmero y coxal de Vilafamés, de acuerdo con la fauna de micromamíferos, se encuadra sin dificultades en el Pleistoceno superior (aun cuando pudieran existir problemas de contaminación paleontológica). Por otro lado, existen pocas dudas sobre la posición estratigráfica de la falange de Cueva Victoria. Según comunicación oral de su descubridor, Joan Pons, esta pieza fue extraída del relleno cárstico denominado Cueva Victoria-Ia, cuya fauna de roedores es sensiblemente idéntica a la del resto de depósitos de la zona. Por otra parte, la asociación de macromamíferos de Cueva Victoria (rinocerótidos, ciervos, carnívoros) parecen indicar una edad anterior a Venta Micena. El reciente descubrimiento de un macaco gigante en el citado yacimiento (género *Paradolichopithecus*, S. Moyà *in verbis*) parece confirmar este último punto de vista. A su vez, Venta Micena no sería exactamente paralelizable a Ubeidiya, ya que su microfauna indica una edad algo anterior. Dejando de lado estas cuestiones menores, hay que alabar la actitud de prudencia que el autor mantiene hacia lo que califica como "una de las fases más obscuras de la evolución humana". Así, en Europa, por ejemplo, se pueden contar con los dedos de la mano el número de mandíbulas atribuibles a esta fase "pre-neandertaliense", en un lapso temporal que abarca más de 700.000 años.

Al estadio de *Homo erectus* está dedicado el trabajo de Wu Rukang y Lin Shenglong sobre la cavidad de Zhokoudian (más conocida por Chou-koutien). Este depósito, al que se asocian nombres tan ilustres como Davidson Black y Teilhard de Chardin, constituye un documento único sobre la evolución física y cultural del "Hombre de Pekín" (antaño denominado *Sinanthropus pekinensis*), entre −460.000 y −230.000 años. El artículo tiene un sesgo muy diferente de los anteriores, centrándose en aspectos generales ligados a la evolución cultural y ambiental, más que en las características físicas de *Homo erectus* en China. Aparte de la impresionante documentación paleoecológica y arqueológica de Zhokoudian, la importancia de esta localidad

queda patente por el alto número de restos craneales completos recogidos.

Finalmente, en lo que hace a este primer apartado, dos de los capítulos seleccionados se refieren a la presencia indirecta, mediante pruebas arqueológicas, de *Homo erectus* en lugares tan alejados entre sí como el sur de la Península Ibérica (A. Ruiz-Bustos: "El yacimiento paleontológico de Cúllar de Baza") y Asia central (R. S. Davis, V. A. Ranov y A. E. Dodonov: "El hombre primitivo en el Asia central soviética"). En concreto, este último trabajo tiene el interés de romper la imagen forzosamente sesgada que tenemos de *Homo erectus* como un poblador de las zonas tropicales o subtropicales del globo: por el contrario, su área de distribución durante el Pleistoceno medio debió extenderse hasta las regiones semiáridas y frías de Asia.

La segunda parte de la obra, correspondiente al estadio *Australopithecus* y su epígono más inmediato, *Homo habilis*, se inicia con el trabajo de A. Walker y R. E. F. Leakey sobre los homínidos del Turkana oriental. El capítulo se inicia con una referencia al cuadro cronológico y tafonómico de los yacimientos de la citada región. Walker y Leakey desgranan a lo largo del trabajo el rosario de notabilísimos hallazgos que se han producido en los últimos veinte años: los cráneos ER-1470 y ER-1813, atribuidos a *Homo habilis*, ER-406 y ER-732 (*Australopithecus boisei*), ER-3733 y ER-3883 (*Homo erectus*), entre otros. De nuevo aquí es de agradecer la prudencia de los autores al evaluar la significación sistemática y evolutiva de los restos del Turkana oriental. Sabedores de los problemas de variabilidad poblacional y dimorfismo sexual existentes en el caso de los primates superiores, Walker y Leakey no excluyen ninguna posibilidad a la hora de emparejar los tres morfotipos que aparecen en esa región (que podríamos catalogar como robusto, grácil arcaico y grácil evolucionado). Después de examinar cada uno de ellos, se llega al modelo comúnmente aceptado durante la última década y que postula la existencia de dos líneas diferentes de homínidos, robusto y grácil, esta última originando *in situ* a los primeros *Homo erectus* (a través de la controvertida especie *Homo habilis*).

Dos pequeñas notas, referentes a recientes hallazgos en Olduvai y Turkana, complementan y actualizan las conclusiones de este último trabajo. Se trata, por una parte, del hallazgo en 1986 del cráneo WT-17000; corresponde a una forma robusta que combina, a la vez, caracteres derivados propios de *Australopithecus boisei* junto con otros claramente primitivos que sugieren una relación directa con *Australopithecus afarensis*. La significación de este resto radica en que excluye una relación directa con las formas sudafricanas de *Australopithecus*. Así pues, a partir de *A. afarensis*, se generarían dos líneas vicarias de australopitecos robustos: de un lado, la oriental (WT-17000, *A. boisei*) y, de otro, la meridional (*A. robustus*). Si se considera a todo el conjunto como monofilético, la resurrección del taxón *Paranthropus* que algunos autores (como el propio Aguirre) proponen parece justificada.

La segunda nota hace referencia al hallazgo por parte del siempre afortunado Johanson de un ejemplar de *Homo habilis* incluyendo fragmentos del cráneo y, lo que es más significativo, parte de su esqueleto postcraneal. Lo interesante en este caso es que, a pesar de su relativamente alta capacidad craneana, el esqueleto de *H. habilis* muestra unas proporciones más parecidas a *Australopithecus* que a *Homo erectus*. En cierta medida, éste era un resultado predecible, ya que muchos de los caracteres de los proclamados *Homo habilis* (ER-1470), OH-24, ER-1813) encajan mejor dentro de una morfología *Australopithecus* que dentro de una morfología de tipo *Homo* (si exceptuamos el tamaño de la dentición y el volumen cefálico). En cualquier caso, la lectura de este conjunto de trabajos arroja un dato evidente: *habilis* fue una cronoespecie de corta duración (no más de 400.000 años) y el tránsito hacia *Homo erectus* debió ser muy rápido. Todo lo cual parece confirmar un modelo puntuado y vicariante para la evolución de los primeros homínidos en Africa.

Finalmente, mencionaremos en este apartado el trabajo del malogrado Glynn Isaac sobre las posibles pautas de comportamiento de los homínidos protohumanos que se deducen de los hallazgos arqueológicos del Turkana oriental. Isaac pasa revista a las diferencias de comportamiento existentes entre nuestros parientes más próximos, los chimpancés, y las sociedades todavía existentes de cazadores-recolectores. Para este autor, nuestros antepasados del Turkana oriental serían merodeadores de los márgenes fluviales y lacustres cuyo comportamiento difería del de los actuales chimpancés al menos en dos aspectos. De un lado, las presas, una vez despiezadas, no serían devoradas *in situ*, sino que serían transportadas a otro lugar. De otro, existiría una auténtica división del trabajo entre machos cazadores y hembras recolectoras que obligaría a compartir el alimento. Estos dos aspectos de transporte de excedentes y reparto posterior serían fundamentales para comprender el curso inicial de la evolución humana. En esta situación, los útiles permitirían la caza de presas de talla grande (superior a los homínidos) y, sobre todo, su despiece posterior.

Por lo que hace al tercer y último apartado de la obra ("Ir a pie y pararse firme. Antecedentes de los homínidos"), el artículo de entrada, a cargo de S. L. Washburn, constituye una puesta a punto sobre los principales problemas con que se encontraba la paleontología humana en el momento de su publicación (1978). Resulta altamente didáctica y clara su comparación con la magnífica síntesis que, sobre una temática similar, publicó D. Pilbeam años después (1984) y que también se incluye en el volumen. Su contraste nos proporciona una adecuada idea de en qué medida nuestra concepción de la evolución humana ha cambiado en los últimos diez años. Otros trabajos en esta sección, se refieren indirectamente al tema de la adquisición del bipedismo. Se trata, por una parte, del artículo de Richard L. Hay y Mary D. Leakey sobre la famosa pista de Laetoli, cuyas pisadas, impresas en una delgada capa de toba volcánica, nos demuestran que hace más de tres millones de años ya existían homínidos bípedos en esta región de Africa. El segundo trabajo, a cargo de D. Campillo, se refiere a la paleopatología de la columna vertebral en poblaciones prehistóricas e ilustra una vez más el hecho de que la evolución raramente llega a soluciones perfectas: la postura erecta, después de 4 millones de años de historia, todavía produce problemas mecánicos en nuestra columna vertebral. El capítulo final de M. P. Ghiglieri sobre la ecología social de los chimpancés proporciona datos de gran interés sobre el efecto fusión-fisión en este grupo.

Estamos, pues, ante una estimulante síntesis cuyas diversas aportaciones, entre 1978 y la actualidad, cubren un excitante período de cambio conceptual en nuestro esquema de la evolución humana, desde el desbancamiento de *Ramapithecus* como posible ancestro de los homínidos africanos hasta los recientes descubrimientos en el Turkana oriental y en la garganta de Olduvai. En definitiva, esta obra, profusamente documentada e ilustrada, puede servir tanto al lector no especializado como al estudiante para revisar, con datos de primera mano, un apasionante capítulo de nuestra historia como especie. (J. A.)

Problemas de Bioquímica, por J. Cárdenas, J. M. Vega, E. Fernández, F. Galván y A. J. Márquez. Editorial Alhambra; Madrid, 1988.

El libro, que se inscribe en una colección de textos auxiliares de clases prácticas, consta de diez capítulos y un apéndice con un total de 490 páginas. En líneas generales cada capítulo se divide en cuatro apartados: conceptos básicos, problemas resueltos, problemas propuestos y bibliografía. La presentación es atractiva, con papel de buena calidad, y la impresión es limpia, con una adecuada distribución de los distintos subapartados.

Un libro de este tipo siempre es recibido con alegría por todos los que nos dedicamos a la docencia de la bioquímica en la universidad. Todos sabemos lo importante que es disponer no sólo de un buen fondo de enunciados de problemas, sino, sobre todo, de su resolución realizada de una manera sencilla y sistemática. Si a ello se añade que los problemas van precedidos de una introducción básica que permite rápidamente repasar los conceptos requeridos para su resolución, la obra alcanza un gran valor didáctico. En este sentido cabe esperar que ello sea de gran utilidad para los alumnos.

En cuanto al contenido concreto de la obra y, al lado de los innumerables aciertos que ya han quedado reflejados anteriormente, hay que dejar constancia de algunos defectos patentes que limitan en cierto modo la utilidad del libro. Así, por ejemplo, si bien en la introducción y en la contraportada se indica que "la materia que aborda corresponde a las unidades temáticas clásicas en un curso de Bioquímica General en la Universidad española: *p*H, tampones, Bioquímica estructural, Enzimología, Bioenergética, Metabolismo y Biología molecular", lo cierto es que en los diez capítulos del libro sólo se tratan las propiedades y estructuras de las proteínas, la estructura de los ácidos nucleicos y la actividad y cinética enzimáticas. No sé si los autores tienen *in mente* la elaboración de otro u otros volúmenes en los que se traten las cuestiones de bioenergética, estructura de glúcidos y lípidos, metabolismo, regulación, etcétera. Está claro, por tanto, que el contenido del libro sólo alcanza algunos aspectos de la bioquímica, quizá los más exigentes desde el punto de vista cuantitativo, pero no los únicos importantes.

Ya se ha comentado que la organización del libro es clara y didáctica en líneas generales. No obstante, en algunos momentos parece que no se haya tenido muy en cuenta la evolución metodológica realmente empleada en los laboratorios. Así, parece muy difícil que algún alumno se haya de enfrentar, en la práctica, con la realización de un análisis de la frecuencia de bases por el método del vecino más próximo. Lo mismo se podría decir acerca de la determinación de la masa molecular de una proteína midiendo las variaciones de la presión osmótica o, incluso, por medio de la ultracentrífuga, a menos que el laboratorio en el que trabaje disponga de dicho aparato y, en cierta manera, se tenga que justificar su adquisición o la existencia del mismo. Por lo que se refiere a los métodos habitualmente más utilizados, hubiera sido interesante indicar que existen algunos comportamientos anómalos, especialmente en el método de la gel-filtración, no atribuibles únicamente a la forma asimétrica de la molécula, sino a interacciones inespecíficas con el gel. En este aspecto de la determinación de masas moleculares creo que es conveniente señalar, y ello es importante tratándose de un libro de problemas, el poco rigor en la utilización indistinta de las expresiones peso molecular y *Mr*.

Respecto a la cinética enzimática creo que hubiese valido la pena ampliar un poco el análisis del comportamiento frente al *p*H, en especial, la distinción entre los valores de *p*K correspondientes al enzima libre (log V/Km *versus* *p*H) y los correspondientes al complejo enzima-sustrato (log V *versus* *p*H). De la primera gráfica no hay ni mención en el texto. También habría sido interesante hablar de los efectos, y forma de analizarlos, que sobre la actividad enzimática en función del *p*H son debidos tanto a la naturaleza de los tampones como a los sustratos que contienen grupos ionizables.

En resumen, hay que celebrar la aparición de un nuevo libro de problemas de bioquímica con una estructuración muy adecuada para los objetivos planteados y en los que algunas de las omisiones son de fácil solución de cara a nuevas ediciones. Para mí, la única objeción realmente importante es al título demasiado general del libro, ya que sólo cubre una parte muy concreta de la bioquímica. (C. M. C.)

Science in the early Roman empire: Pliny the elder, his sources and influence. Dirigido por Roger French y Franck Greenaway. Croom Helm; Londres, 1986.

Es cortesía que el lector espera del crítico le exponga éste el estado de la cuestión, en nuestra área idiomática, del tema reseñado. ¿Qué sabemos aquí de Plinio? Por lo publicado, apenas nada más que lo que traen las enciclopedias, fusilado de otras foráneas. No existe una edición crítica castellana de su *Historia Natural*, como sí disponen de la suya alemanes, ingleses y franceses. Y algunas incursiones de supuestos expertos más valiera no haberlas realizado.

En un catálogo de libros de historia natural, editado dos años más tarde que el volumen que encabeza este comentario, los autores se sintieron en la obligación de resumir en la introducción la historia del concepto de historia natural. Tras una incorrecta transcripción de las vocales griegas con espíritu áspero, la del artículo femenino y del término *historía*, que nadie pide, pero que si se dan han de darse bien, y la omisión de la obra de Teofrasto, en la que se espejaron las "historias naturales", hablan de Plinio con voz de diccionario. Se nos dibuja allí un estereotipo roto por la crítica hace tiempo, resumida en *Plinio e la natura* (Como, 1982) y, muy especialmente, esta *Science in the early Roman empire*. Por no entrar en matices de fecha de nacimiento, lugar de fallecimiento, etcétera, al abordar la causa de su muerte –sobre la que existe una interesante controversia–, los autores del catálogo atribuyen la descripción de la misma a "algunos de sus biógrafos", la que, sabido es, débese a su sobrino Plinio el Joven, en carta a Tácito: "Petis, ut tibi avunculi mei exitum scribam... Erat Miseni classemque imperio praesens regebat...". "Me ruegas que escriba sobre la muerte de mi tío... Estaba en Miseno, al frente de la escuadra..." (*Epistularium libri decem*, VI, 16). El resto de la introducción sobre el desenvolvimiento de la historia natural no aportará al lector lo que no pueda encontrar en cualquier compendio de urgencia, aunque a veces sí: halla lo que no debe decirse. "En el Renacimiento... se fundan los primeros jardines botánicos[6]..." Nota que reza: "[6]Montpellier (1593), Oxford (1621), Jardin du Roi de París (1635)...". La historia de los jardines botánicos en el Renacimiento comienza mucho antes. Por apoyarme en los manuales: "Gardens had been founded in Pisa, Padua and Florence by 1546, and the next twenty years saw their establishment in Ferrara, Sassari, Bologna and many other places in Italy" (A. G. Morton: *History of Botanical Science*, p. 121). No vale la pena seguir. Ni entrar en la confección del catálogo. La verdad es que ni siquiera hubiéramos mentado la alusión a Plinio, si el repertorio no lo hubiera publicado el Consejo Superior de Investigaciones Científicas. *Et de hoc satis*.

Vayamos a lo serio. Coordinada por F. Greenway, que se ocupa de la química en Plinio, y R. K. French, que estudia la influencia del romano en la medicina del Renacimiento, asistimos a una investigación exhaustiva de varios aspectos de su obra: relación con la medicina romana, farmacia, botánica, zoología, metalurgia y mineralogía, química y astronomía.

Especialmente ilustrativos, por lo que tienen de paradigmáticos, el capítulo segundo ("El grupo traductor alemán de Plinio") y el tercero ("La estructura de la *Historia Natural* de Plinio el Viejo"). El "Projekgruppe Plinius" reúne, desde 1976, un equipo de profesores de distintas disciplinas para una nueva versión germana de la obra máxima del romano. Hasta ahora, las ediciones críticas pertenecían al dominio de los filólogos o naturalistas con preparación humanística, que echaban mano de diccionarios etimológicos, gramáticas comparadas y demás adminículos para establecer el significado de un término. Así la "Budé" o la "Loeb". El grupo germano viene trabajando con otro enfoque. Lo sintetizaremos con un ejemplo. ¿Qué es, en Plinio, el "plumbum argentarium"? ¿Plomo plateado? Es decir, ¿la mitad de plata y la otra de plomo? Cuando Plinio habla de "plumbum album" hay coincidencia en admitir que alude al estaño y, cuando se refiere al "plumbum nigrum", al plomo. Tras una prueba de metalurgia con "aes" (unas veces bronce, otras cobre) R. C. A. Röttländer, autor del capítulo segundo, establece que "plumbum argentarium" debe ser el plomo, frente a la opinión establecida que lo suponía estaño.

J. Scarborough, especialista en farmacia romana, con varios trabajos sobre Teofrasto, Dioscórides y el propio Plinio, se explaya en las fuentes posibles donde bebió nuestro autor. En 27 páginas antológicas desmadeja el embrollo de influencias, desde Aristóteles hasta las postrimerías del siglo I de nuestra era, en la línea que J. Stannard había ya esbozado con sumaria claridad, deteniéndose especialmente en la figura capital de Sextius Niger. Reconocimiento que debemos prestar también al trabajo de O. Pedersen sobre algunos puntos de la astronomía de Plinio; salva a éste del infierno de la astrología y se extiende en el valor práctico –el "rústico" que Plinio fue siempre, con la mirada atenta en el cielo– de sus observaciones astronómicas.

De la sombra alargada de Plinio en la Edad Media y el Renacimiento se ocupa B. S. Eastwood, quien prolonga en esos períodos el estudio astronómico de Pedersen (ambos trabajos acompañados de aparato gráfico), y R. K. French, quien aborda las disputas plinianas y esquematiza la fuerza, apenas conocida, de los llamados helenistas: grupo de científicos y filósofos que, venidos de Bizancio, más o menos en torno al concilio de Ferrara, recalaron en esta ciudad y dieron un gran impulso al pensamiento griego. French se apoya en los estudios de Fridolf Kudlien en la identificación del anónimo contra el que disputa Leoniceno, si bien convendría haber anotado que la expresión "Plinii patronus" se la aplicó ya el de Vicenza a Angelo Poliziano.

Plinio es, sin duda, uno de los clásicos que está demandando la atención de los historiadores de la ciencia española. Por volver al grupo alemán como punto de referencia, no se olvide que describió con especial pormenor el beneficio de los minerales en nuestro país, uno de los suministradores de oro al imperio. (L. A.)

BIBLIOGRAPHIA CHEMICA HISPANICA, 1482-1950, por Eugenio Portela Marco y Amparo Soler Sáiz. Instituto de Estudios Documentales e Históricos sobre la Ciencia; Universidad de Valencia.-CSIC, 1987.

Es éste el segundo, y único hasta ahora aparecido, de una obra que se ha proyectado en cuatro volúmenes: libros y folletos, 1482-1800; libros y folletos, 1801-1900; libros y folletos, 1901-1950, más índices; revistas. Se inscribe en una colección que ha dado ya a luz pública distintos repertorios que constituyen una herramienta indispensable para los profesionales de la historia de la ciencia y para cualquiera que se adentre en ese campo.

Se sigue un orden alfabético de autores, tras cuyo nombre se enuncia el título de la obra, ciudad e impresor, año de edición, biblioteca pública en que se halla y formato. Cuando el libro o folleto es anónimo, la primera palabra (en negritas) encabeza la entrada. Otros datos incluidos son la signatura de la biblioteca, el nombre del traductor, los repertorios de donde se extrae y otras anotaciones de interés. En la introducción se indica el criterio seguido para delimitar lo que entienden los autores por "química".

Ni siquiera los medios informáticos de que dispone el Instituto de Estudios Documentales de Valencia, lugar de trabajo de los autores, excluyen las posibles omisiones, que vienen a ser, en este tipo de obras, como las erratas en los libros mejor editados. No veo, por ejemplo, el *Nuevo Diccionario de Química* (2 vols.) de Emilio Bouant, traducido y adaptado por Manjarrés y Tremols, ni la tercera edición del Wurtz traducido por Almera (Barcelona, 1888). Y en este capítulo de omisiones, convendría rehacer las dos entradas relativas a Berzelius e incorporar el "*Tratado de química mineral, vegetal y animal*" por... segunda edición francesa, traducida... por M. M. Esslinger y Hoeffer de la quinta edición...", en 15 volúmenes.

Otra fuente de error yace en la propia copia de los datos. Sin salirnos de obras que ejercieron especial influencia, las *Lecciones de química elementar* (no "elemental", y no se repute mero arcaísmo) de Jean Girardin son "traducidas de la *segunda edición francesa* y adicionadas" (donde la cursiva indica texto omitido). ¿Por qué esperar a la séptima edición del manual de Rico-Santisteban para señalar el número de grabados intercalados en el texto? ¿Por qué no anotar el número de partes, editadas por separado, del *Traité des poisons* de Orfila, amén de repasar el texto genuino del título?

Alguien podría pedir cierta coherencia en algunas fichas. Verbigracia, el título de la primera obra registrada de Justo Villanueva debe rezar exactamente: *Estracto de las lecciones de Química aplicada a la Agricultura esplicadas*... Herederos de la Viuda Pla", paralelo al primero de la página siguiente (250). De lo contrario, da la impresión de que uno se ha cotejado y el otro no. Abundando en ese apremio de coherencia, el primero de los *Estracto* citados debería haberse paginado así: 8 + 288.

Algunas equivocaciones son travesuras de los duendes de la linotipia: los *Elementos de química médica*, de Orfila es una obra traducida; el pie de imprenta del libro de su émulo Mata, *Compendio de Toxicología*, es *Merás* y Suárez, no Merino y Suárez; por cierto, el número de páginas del mismo es de 499 + Advertencia. ¿Están en cuarto o en octavo los *Elementos de farmacia* (1805) de Carbonell Bravo?

Haría muy mal el lector si creyera que se halla ante una obra precipitada y carente de rigor. Si hemos destacado algunas lagunas someras, subsanables fácilmente en una próxima edición, ha sido para ponderar la dificultad de ese trabajo benedictino, incluso realizado con la profesionalidad del centro valentino, cuyas líneas de trabajo acaba de resumir su director en *Nuncius*, importante revista de historia de la ciencia. Otros repertorios más recientes, publicado alguno por nuestro instituto oficial de investigación, producen grima y sonrojo ajeno. (L. A.)

Apuntes

La anemia falciforme es la primera enfermedad congénita de cuya existencia se enteran los escolares. Con su ejemplo aprenden que basta una mutación en un par de bases del gen de la betaglobina (trueque de adenina por timina) para que se produzca un cambio en los aminoácidos de la proteína (sustitución de ácido glutámico por valina), sustitución que puede llevar, en determinadas circunstancias, a un desenlace fatal. Pero el conocimiento del proceso bioquímico implicado no ha supuesto el dominio de la terapia génica, y se ha convertido en piedra de toque para valorar el poder sanador de la nueva genética.

Sigamos con la mutación. Dice la teoría clásica de la evolución molecular: la tasa de mutación puntual es uniforme en todo el genoma de un organismo y la modificación de la tasa de sustituciones nucleotídicas en regiones de ADN refleja las tensiones selectivas diferenciadoras. No parece que vayan a durar mucho esas afirmaciones si se corrobora la elevada variabilidad de la tasa de mutación observada ya en determinados segmentos del genoma de mamíferos.

La fusión de cuerpos planetarios pequeños en el comienzo del Sistema Solar encontró en la desintegración radiactiva del aluminio-26 en magnesio-26 la fuente de calor necesaria. Tal hipótesis se corrobora en la primera observación de magnesio de origen radiactivo presente en un cóndrulo de la condrita de Semarkona. Hallazgo que, a su vez, sugiere una nueva hipótesis: la acreción y diferenciación planetarias hubo de comenzar en una escala temporal comparable con la vida media del aluminio-26.

De las tres hojas embrionarias de los vertebrados, el mesodermo es la última que aparece, y viene inducida en el ectodermo por señales que proceden del endodermo subyacente. Se sospecha que esas señales inductivas podrían condicionar las pautas configuradoras de los organismos vertebrados. ¿En qué consisten esas señales? Se ha empezado a hablar de ciertas moléculas implicadas: los factores de crecimiento, entre ellos el de los fibroblastos.

Las innovaciones técnicas han abierto siempre nuevas perspectivas a las innovaciones teóricas. Eso ocurrió con la invención de la resonancia magnética nuclear, a mediados de los cuarenta. En su postrer refinamiento, la resonancia magnética nuclear de rotación doble, la muestra a observar gira rápidamente, en un campo magnético, en torno a dos ejes al mismo tiempo. Ensayada con éxito en la identificación de la estructura de los núcleos de sodio-23 en mezclas de sulfato y oxalato de sodio, se espera que geólogos, ingenieros de materiales y estudiosos de catalizadores, adsorbentes y superconductores de altas temperaturas encuentren en esa técnica una poderosa arma de trabajo.

Por lo menos en las plantas la hembra sí es el sexo débil. La reproducción sexual en ese reíno implica un altísimo coste energético, que debe invertirse en la floración y fructificación. La planta femenina gasta mucho más que la masculina, aunque ésta pueda superarle a la hora de echar flores. Debido al mayor sacrificio exigido por la fructificación, la planta femenina suele mostrar una menor tasa de desarrollo, florece con menos frecuencia y vive también menos.

En las hembras animales, por contra, la mayor inversión se da en el período de amamantamiento. Más aún, comparada con la gastada en esta fase, la energía dedicada a la gestación es mínima si se trata de poblaciones con recursos alimentarios limitados. Por ello mueren tantos juveniles en esas circunstancias, obligada la madre a concentrarse en la supervivencia de los que realmente puede amamantar.

El amoniaco es tóxico para todos los vertebrados. Puede convertirse en urea, menos dañina para la supervivencia. Pero ese proceso de conversión es muy costoso desde el punto de vista metabólico y sólo se encuentra en las especies terrestres que no pueden excretar fácilmente amoniaco y en los peces de agua salada que recurren a la urea como regulador osmótico. Los peces teleósteos de agua dulce han suprimido el ciclo metabólico de producción de urea. Pero la única especie teleóstea que ha logrado sobrevivir en un lago intensamente alcalino, el pez *Oreochromis alcalicus grahami*, se las ha maravillado para desarrollar las enzimas hepáticas necesarias en el ciclo de la urea.

Seguiremos explorando los campos del conocimiento



LA AVENTURA ESPACIAL SOVIETICA, por Peter M. Banks y Sally K. Ride
Desde 1971, los cosmonautas han permanecido a bordo de naves espaciales soviéticas más de 5600 días en total. Sin embargo, los vuelos tripulados representan sólo una pequeña parte del ambicioso programa espacial de la Unión Soviética.

GIGANTESCAS CATARATAS OCEANICAS, por John A. Whitehead
Las mayores cascadas de agua en la Tierra no son las célebres cataratas, sino corrientes invisibles en las profundidades oceánicas, impulsadas por diferencias de temperatura o de salinidad.

GENES DE LA VISION DEL COLOR, por Jeremy Nathaus
Acabados de aislar, los genes que determinan las proteínas detectoras del color en los seres humanos muestran el camino de la evolución de la visión cromática normal y apuntalan la investigación sobre la ceguera a los colores.

LOS NUEVOS SUPERCONDUCTORES, por Alan M. Wolsky, Robert F. Giese y Edward J. Daniels
La ciencia de los materiales tendrá que desarrollarse más si quiere satisfacer las esperanzas puestas en los superconductores de altas temperaturas.

DIENTES Y PREHISTORIA EN ASIA, por Christy G. Tuerner II
Sutiles diferencias en la estructura dentaria nos permiten reconstruir grandes movimientos migratorios en el Nuevo Mundo, Asia oriental y cuenca del Pacífico.

METABOLISMO FOTOSINTETICO DEL CARBONO EN LAS PLANTAS SUPERIORES, por Daniel H. González, Alberto A. Iglesias, Florencio E. Podestá y Carlos S. Andreo
Algunas especies de plantas superiores han desarrollado rutas auxiliares para la asimilación del carbono. Estas permiten aumentar la productividad vegetal y el desarrollo de la planta en condiciones ambientales desfavorables.

DEL CANTO DE LOS PAJAROS A LA NEUROGENESIS, por Fernando Nottenbohn
El estudio de los centros de control del canto en el cerebro de los canarios revela que en el cerebro adulto nacen células nerviosas y que éstas pueden sustituir a células más antiguas. Esta forma de neurogénesis puede ser la clave de la autorreparación del cerebro.

LOS EFECTOS QUIMICOS DE LOS ULTRASONIDOS, por Kenneth S. Suslick
Cuando las ondas ultrasónicas intensas atraviesan los líquidos, se forman pequeñas cavidades que aumentan de tamaño y finalmente implosionan, produciendo una enorme cantidad de calor. Estas condiciones extremas proporcionan un medio químico de características singulares.

